Impact of Imaging and Distance Perception in VR Immersive Visual Experience

GIUSEPPE MORANA

PRINCIPAL SUPERVISOR: DR. SALVATORE LIVATINO
SECONDARY SUPERVISOR: DR. ALESSIO MALIZIA

Thesis submitted to the University of Hertfordshire in partial fulfilment of the requirements of the degree of Doctor of Philosophy (PhD)

United Kingdom
August 2023

Abstract

Virtual reality (VR) headsets have evolved to include unprecedented viewing quality. Meanwhile, they have become lightweight, wireless, and low-cost, which has opened to new applications and a much wider audience. VR headsets can now provide users with greater understanding of events and accuracy of observation, making decision-making faster and more effective. However, the spread of immersive technologies has shown a slow take-up, with the adoption of virtual reality limited to a few applications, typically related to entertainment. This reluctance appears to be due to the often-necessary change of operating paradigm and some scepticism towards the "VR advantage". The need therefore arises to evaluate the contribution that a VR system can make to user performance, for example to monitoring and decision-making. This will help system designers understand when immersive technologies can be proposed to replace or complement standard display systems such as a desktop monitor.

In parallel to the VR headsets evolution there has been that of 360 cameras, which are now capable to instantly acquire photographs and videos in stereoscopic 3D (S3D) modality, with very high resolutions. 360° images are innately suited to VR headsets, where the captured view can be observed and explored through the natural rotation of the head. Acquired views can even be experienced and navigated from the inside as they are captured.

The combination of omnidirectional images and VR headsets has opened to a new way of creating immersive visual representations. We call it: *photo-based VR*. This represents a new methodology that combines traditional model-based rendering with high-quality omnidirectional texture-mapping. Photo-based VR is particularly suitable for applications related to remote visits and realistic scene reconstruction, useful for monitoring and surveillance systems, control panels and operator training.

The presented PhD study investigates the potential of photo-based VR representations. It starts by evaluating the role of immersion and user's performance in today's graphical visual experience, to then use it as a reference to develop and evaluate new photo-based VR solutions. With the current literature on photo-based VR experience and associated user performance being very limited, this study builds new knowledge from the proposed assessments.

We conduct five user studies on a few representative applications examining how visual representations can be affected by system factors (camera and display related) and how it can influence human factors (such as realism, presence, and emotions). Particular attention is paid to realistic depth perception, to support which we develop target solutions for photo-based VR. They are intended to provide users with a correct perception of space dimension and objects size. We call it: *true-dimensional visualization*.

The presented work contributes to unexplored fields including photo-based VR and true-dimensional visualization, offering immersive system designers a thorough comprehension

of the benefits, potential, and type of applications in which these new methods can make the difference.

This thesis manuscript and its findings have been partly presented in scientific publications. In particular, five conference papers on Springer and the IEEE symposia, [1], [2], [3], [4], [5], and one journal article in an IEEE periodical [6], have been published.

Declaration Statement

I certify that the work submitted is my own and that any material derived or quoted from the published or unpublished work of other persons has been duly acknowledged (ref. UPR AS/C/6.1, Appendix I, Section 2 – Section on cheating and plagiarism).

Student Full Name: Giuseppe Morana

Student Registration Number: 17070885.

Signed: Giuseppe Morana

Date: 03/08/2023

Acknowledgements

I dedicate this thesis to all the people that have supported me with their love and have never stopped believing in me since the very beginning of this long challenging journey.

I feel extremely lucky to have worked at the University of Hertfordshire's Virtual Reality and Robotics Lab, where I met friendly and inspiring fellow students, who were very supportive of me with their technical knowledge and friendliness. I would especially like to thank Phil and Yasir.

I would like to thank my supervisor Dr. Salvatore Livatino for being a patient and an educational supervisor, for his friendly and scientific guidance, and for having supported me throughout the time I was at the University of Hertfordshire.

I would also like to thank those who have helped me in various ways at specific moments in my scientific journey. In particular, Antonino, Nicoletta, Luca, Francesco G., Francesco R., Giovanni, Alessio and Abid.

A special thanks goes to my family, who managed to be present every day for the last years despite the long physical distance, helping me morally to bear all the obstacles of my new adventure in the UK. I would like to thank my mother for her priceless daily support, and for giving me life and all her love. I would also like to thank my sister Carmen, her husband Giovanni and my two fantastic nieces Arianna and Elena, for their loving support and comforting presence. I would like to thank my father who, even from up there, continues to guide me and my future towards great, profitable and surprising experiences. I'm sure he would have been very proud of my result.

Many thanks to my second UK family, Salvo, Kasia and their two wonderful children Alex and Emily, who have warmly and kindly welcomed me into their life.

A special mention goes to my best friend Riccardo for the invaluable evenings spent in discussion at The Boot. I'd also like to mention the amazing flatmates I've lived with over these years: Paula, Sally, Nick and Peter.

Finally, I want to especially thank my girlfriend Tanya, who is the person who supported me the most during my PhD years, giving me all her effort, energy and love to help me make this dream come true. You have all my gratitude, and my love.

This would not have been possible without you. I love you all!

Giuseppe Morana

Contents

1. Introduction

This chapter briefly introduces the main topic of the PhD project, which concerns the study and evaluation of the immersive visual experience and the development of new solutions to enhance some aspects of the visual experience. The focus is on the use of photographs and improving the perception of depth.

The chapter therefore begins by introducing: the main aspects of the immersive visual experience, the technology that provides this experience, the potential and limitations of the use of photographs in VR, and the challenges associated with a realistic perception of depth.

The arguments introduced form the basis for then outlining the proposed investigation and briefly explaining the reasons supporting this research. Finally, the content of each chapter of the dissertation is briefly mentioned.

1.1 Immersive Visual Experience and Technology

In our rapidly advancing digital age, the pursuit of immersive visual experiences has captivated the minds and hearts of individuals around the world. Whether it is through virtual reality (VR), augmented reality (AR), or other groundbreaking technologies, the possibility of extending our physical reality and has brought us new dimensions and application possibilities.

Immersive visual experiences bring a transformative power, which has the potential of reshaping various aspects of our lives, from entertainment to action training.

What affects users' actions when observing remote events on a visual display, is their visual comprehension of the observed scene. These impacts monitoring effectiveness and the decision-making processes. The role that immersive technology plays in terms of visual experience is then of paramount importance and worth being studied, assessed, and developed.

When studying the immersive visual experience, a key role is played by the system's provided sense of immersion. This profoundly affects user' sensation, as it makes users feel present in the observed virtual world. The resulting sense of telepresence greatly influences user's behaviour, and his/her performance.

Immersive visual experiences and VR technology have become synonymous with a new era of digital engagement. By merging advanced graphics, motion tracking, and sensory feedback, VR creates a simulated environment that surrounds and engulfs users, transporting them to extraordinary realms and enabling unparalleled levels of immersion. The symbiotic relationship between immersive visual experiences and virtual reality technology can revolutionize various domains, from entertainment and remote exploration to training and teleoperation.

Virtual reality technology serves as the foundation for immersive visual experiences, offering users the ability to step into artificial worlds and interact with them in ways previously unimaginable. With the aid of VR headsets, users are able to perceive three-dimensional environments that respond to their movements and actions, fostering a sense of presence and realism. This technology enables a profound level of immersion, engaging multiple senses and stimulating a deep emotional connection between users and the virtual world.

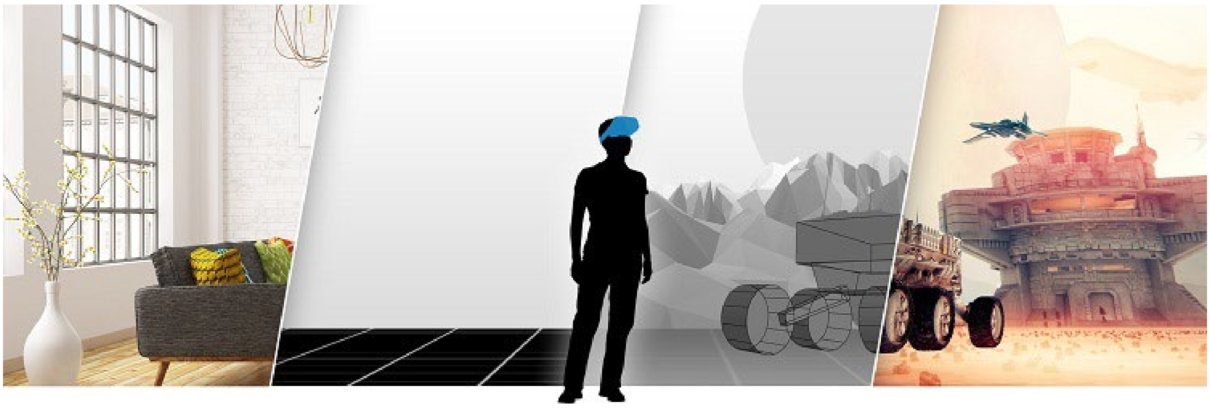


Figure 1 - Virtual Reality Immersive Experiences.

1.2 VR Visual Realism and Depth Perception

The future of immersive visual experiences and virtual reality technology holds immense potential. As technology continues to advance, VR headsets become more affordable, and we can anticipate even more immersive and transformative experiences, extending the applications of immersive visual experiences and VR technology.

VR technologies offer immense potentials in terms of creating realistic and immersive experiences that can transport users to virtual worlds. The potential of VR technologies to create realistic visual experiences is significant. By utilizing high-resolution displays, advanced graphics rendering, and precise tracking systems, VR can provide visually immersive environments that resemble real-world settings.

However, VR realism also faces limitations when it comes to achieving complete realism because of the time-consuming processing and the complexity of rendering photorealistic appearance. One limitation is the resolution and visual fidelity of current VR displays. While modern VR headsets offer impressive resolutions, they may still fall short of the visual acuity and detail found in the real world. This can lead to a slight compromise in the realism of virtual environments.

Texture mapping is then often employed because of its ability of integrating within a photographic image, complex lighting and reflection effects. The use of photographic images can provide a level of photometric interaction (e.g. colour, contrast, definition, vividness) difficult to generate through graphics. This realism allows users to feel a sense of presence and engagement, enhancing the overall experience across various domains and consequently their action performance. When using photographs there are however shortcomings. There is a need to map specific images to specific model surfaces. Furthermore, a photograph represents a specific viewpoint of observation, departing from which visual quality degrades and deformations may appear.

Another relevant element of visual realism is the provided depth sensation a view contains. Depth perception is a crucial aspect of creating realistic experiences in VR. By simulating depth cues such as stereoscopic vision, parallax, and convergence, VR can provide a sense of depth and distance in virtual environments. This allows users to perceive objects and spaces as they would in the physical world, enhancing the feeling of immersion and presence.

Unfortunately, the depth impression that a VR view provides is often different of that intended by the system designer. It is in fact difficult to accurately simulating depth perception cues in VR. While stereoscopic rendering provides a three-dimensional sense of depth, it may not replicate all the depth cues that our visual system relies on in the physical world. Furthermore, parameters such as lens distortion can negatively impact realistic perception of distances.

In case of VR views representing real places, a critical issue is providing users with the same depth impression experienced in the real place. Photographs may include several photometric characteristics of the real place, which increase realism. However, distortions may rise due to image-capture settings, image processing and transmission, visualization process, visual rendering quality, visualization interocular distance (IOD) and display field of view (FOV).

There are then other elements beside realism and depth perception that can affect the VR immersive visual experience. The tracking system and its latency is for example crucial for maintaining the illusion of depth and spatial presence. Any delays or inaccuracies in tracking can cause break of presence or can lead to potentially compromise depth perception and overall realism of the experience.

Despite continuous advancements in VR technologies are being made to improve realism and depth perception, addressing the mentioned limitations for example by understanding their causes is deemed key elements towards improving the fidelity of VR experiences. It can unlock new levels of immersion and depth perception, bringing us closer to the goal of truly realistic virtual worlds. This has great potential towards consideringly widening VR application fields.

Figure 2 - Depth Perception in VR improve truly realistic virtual worlds.

1.3 Proposed Investigation

The core idea for the PhD proposed investigation is to study for VR applications the combined use of *immersion*, *photographic texture* and *true-dimensional visualization*.

The potential, and therefore the interest for this study come from the following facts:

- *Immersion.* VR systems are great for providing immersion to users.
- *Photographic texture*. Image capture and recording are now more than ever available, low-cost and capable of providing high-resolution 360-degree views.
- *True-Dimensional Visualization*. The ultimate and most popular VR system, i.e., the HMD, can favorably support realistic depth perception. Wearing an HMD gives users the possibility to observe the virtual world the way we naturally observe it every-day. The FOV may resembles the one humans naturally perceive and the environment around us can be discovered by turning our head.

Research is needed to best align and exploit immersion, photographic texture and true-dimensional visualization. This means first of all to understand key characteristics, advantages and disadvantages they bring to applications, as well as to study and assess the role of the main involved parameters.

A number of user studies are proposed to assist and inform the development of new solutions. They have the aim to improve performance of the proposed combination, while paying specific attention to imaging and distance perception. The thesis' chapter 3 presents more specifically the core idea, motivation, and development plan. The chapters 4 through 6 describes in detail the research work carried out in the project and its outcomes.

1.4 Thesis Outline

Chapter 1 introduces the main topics of the PhD project briefly reporting on immersive visual experience and VR technology, including their advantages and limitations. It outlines the proposed investigation and chapters content.

Chapter 2 introduces the most relevant background knowledge needed to comprehend the subject area and support the sought objectives. The chapter overviews immersive technologies and the main visualization techniques. It includes description of the related system parameters and human factors.

Chapter 3 presents the core idea and argumentation for the proposed study. It introduces the three main aspects of investigation, which are presented in greater depth in chapters 4 through 6. Chapter 3 also includes the study project development plan.

Chapter 4 specifically addresses immersion and presents two user studies where objective and subjective data have been collected to demonstrates the immersion advantages within specific application contexts. The presented studies involve actions of monitoring and teleoperation as well as user-scene interaction.

Chapter 5 specifically addresses photo-based VR and presents two user studies where data related to several display and human factors have been collected. The presented studies compare the use of two different screens, locations and senses of place. Conclusions and some summarizing key points are also drawn.

Chapter 6 addresses depth perception and distance estimation. It starts with presenting a user study that collected user's depth impressions and tested user's distance-estimation accuracy. It then describes some developed techniques to reduce deformations and improve depth perception in photo-based VR. The last section presents the design and setup of a new procedure to reduce and counterbalance arising viewing distortions.

Chapter 7 draws the project conclusions, summarizes achievements, and identify some aspects for future research.

2. VR Visual Experience

2.1 Computer Graphics Visualization

In this section the concept of computer graphic visualization is introduced. Computer graphics relies on the use of 3D modelling where objects contained in the represented virtual environment modelled geometrically and photometrically, i.e., they are given shape and texture. The object’s texture can be either synthetically generated or provided through a photograph. The latter is typical in what we call photo-based virtual reality.

2.1.1 Modelling and Rendering

Computer graphics modelling and rendering are fundamental processes in the creation of visually compelling digital content. Modelling involves the creation and manipulation of 3D virtual objects and scenes, while rendering focuses on transforming these models into 2D images or animations.

Modelling encompasses various techniques to represent objects accurately and efficiently. Polygonal modelling employs polygons, such as triangles and quadrilaterals, to define an object's shape. Curves and surfaces, like Bezier curves and NURBS enable the creation of complex and smooth shapes. Procedural modelling employs algorithms to automatically generate models, allowing for the creation of intricate and realistic scenes. Volumetric modelling represents objects using voxels or other volumetric data structures, particularly useful in medical imaging and scientific visualization.

Geometric transformations play a vital role in positioning and manipulating objects in 3D space. Translation shifts objects, rotation changes their orientation, scaling alters their size, and shearing distorts objects along specific axes.

Lighting and shading techniques simulate the interaction between light sources, materials, and the viewer's perspective. Illumination models mathematically describe how light interacts with surfaces, while shading calculates the colours of pixels based on the lighting conditions and surface properties. Global illumination techniques, like radiosity, replicate realistic lighting effects by considering indirect lighting and light bounces.

Texture mapping enhances the visual appearance of objects by assigning 2D images (textures) to 3D models. UV mapping coordinates the mapping process, ensuring the textures fit properly onto the model's surface.

Rendering involves the transformation of 3D models into 2D images or animations. Rasterization is a common technique that converts objects into pixels on a 2D grid, while ray tracing simulates the path of light rays to generate highly realistic images. Hybrid approaches combine rasterization and ray tracing for efficient and visually appealing rendering.

Computer graphics modelling and rendering pipelines organize the various stages of the process. The modelling pipeline involves creating and manipulating 3D models, while the rendering pipeline encompasses transforming the models, applying lighting, and shading, and generating the final output. Understanding these concepts and techniques are crucial for creating captivating visual content, whether in the fields of animation, gaming, virtual reality, or scientific visualization. By mastering computer graphics modelling and rendering, one gains the ability to bring imagination to life and communicate ideas.

2.1.2 Mapping Images into Spheres and Cubes

Creating a realistic representation of a spherical object requires mapping 2D image onto a 3D sphere surface.

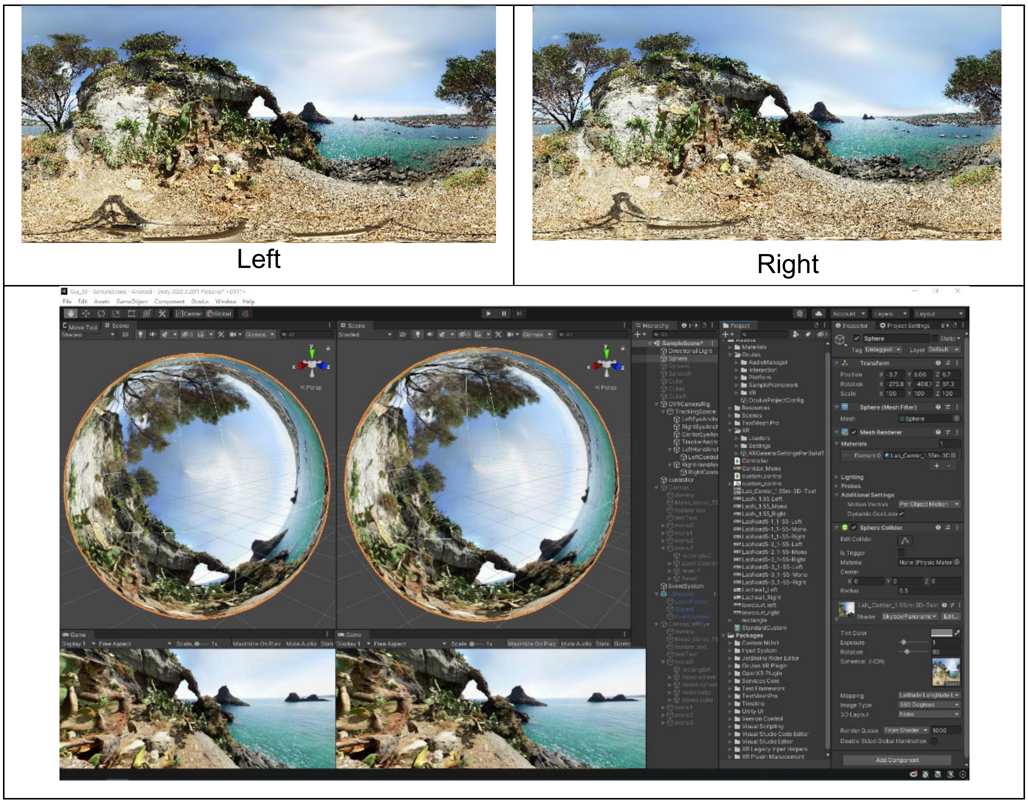


Figure 3 - Mapping Images (Left and Right) into a Speres with Unity Software.

Mapping an image onto a sphere surface in computer graphics is a powerful technique that enables the creation of immersive and realistic visual experiences. It finds applications in various fields such as virtual reality, video games, architectural visualization, and scientific simulations, allowing for the accurate portrayal of spherical objects in a digital environment.

This technique is commonly for tasks such as rendering Earth globes, celestial bodies, or immersive 360-degree panoramas. The process begins by defining a spherical geometry to serve as the surface to map an image. The sphere can be created by dividing a regular polyhedron, such as an icosahedron, into smaller triangular faces. Each vertex on the sphere corresponds to a point in 3D space, and these vertices are assigned texture coordinates to establish the mapping relationship between the sphere and the image.

The next step is to unwrap the sphere onto a 2D plane. This unwrapping process typically involves converting the 3D coordinates of the sphere to a 2D coordinate system while preserving the spherical nature of the object. One common method is to use latitude and longitude values to parameterize the surface of the sphere. The vertices of the sphere are mapped to corresponding points on the 2D plane, and the texture coordinates are generated accordingly.

Once the sphere is unwrapped and the texture coordinates are established, the image can be applied to the surface. The 2D image is projected onto the unwrapped 2D plane using techniques such as texture mapping (UV mapping). The texture coordinates obtained from the unwrapping process map to the corresponding pixels on face or vertices. This process ensures that the image is aligned and stretched to fit the shape of the original sphere.

During rendering, the mapped image is interpolated across the surface of the sphere to ensure a smooth and seamless transition between the pixels. This helps to create a realistic and visually appealing representation of the spherical object. The result is a textured sphere that accurately providing a convincing 3D representation of a spherical object in computer graphics.

Mapping an image onto a cube surface in computer graphics involves projecting a two-dimensional image onto the six faces of a three-dimensional cube. This technique, often referred to as cube mapping, cube texture/ Equirectangular mapping creates realistic textures and environment mapping in 3D rendering.

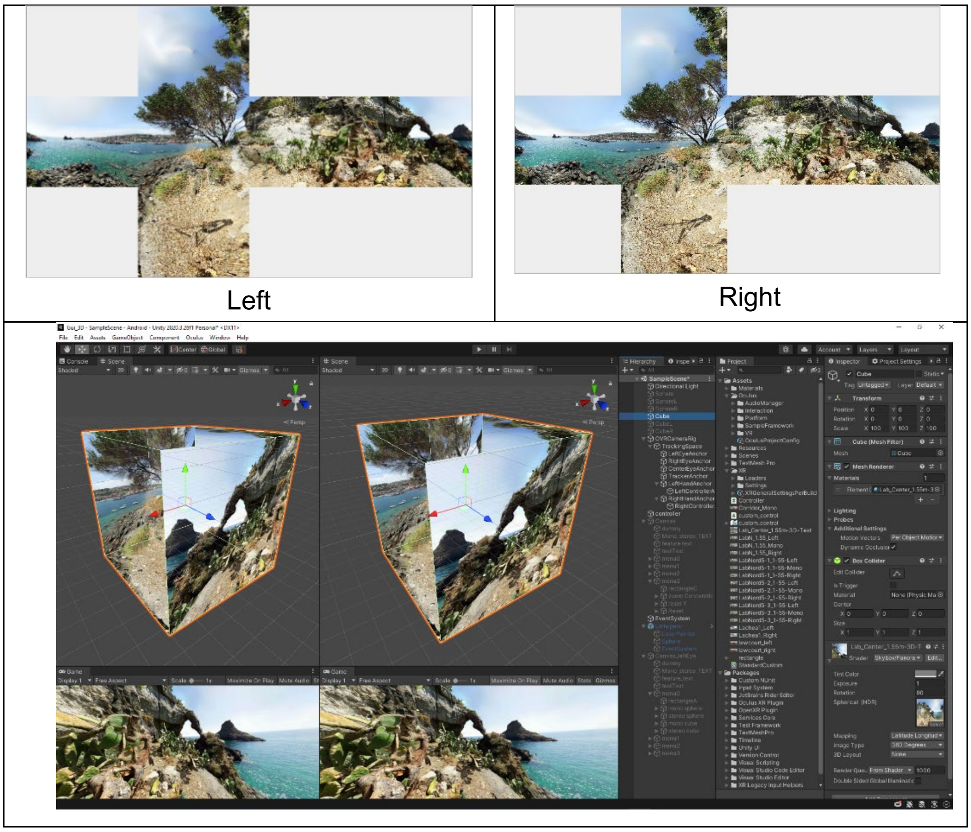


Figure 4 - Mapping Images (Left and Right) into a Cube with Unity Software.

The process begins with selecting or creating an image that will be applied to the cube. This image is typically in a rectangular format, representing the texture or environment that will be mapped onto the cube. The image should be carefully designed to align seamlessly across the cube's faces.

To perform the mapping, the cube is first defined in 3D space with its vertices and faces. Each face of the cube is then unwrapped or flattened into a 2D plane. This process transforms the cube's faces into six individual rectangles, maintaining the relative proportions and angles of the original cube.

The next step involves associating specific coordinates in the 2D image with each vertex of the flattened faces. These texture coordinates are used to determine how the image will be mapped onto the cube. Mapping techniques like UV mapping are commonly used to define the correspondence between the image and the cube's faces.

Once the texture coordinates are assigned, the image is then mapped onto each face of the cube. This is achieved by applying the texture coordinates to the corresponding vertices of each face. The texture is then interpolated across the face, determining how the image is stretched or distorted to fit the surface.

When rendering the cube, each pixel on the surface is assigned a direction vector based on the planes normal. This vector is used to determine the corresponding texture coordinates on the cube's faces. By sampling the image at those coordinates, the color information is applied to the pixel, creating the illusion of a textured or environment-mapped surface.

Mapping an image onto a cube surface is a versatile technique used in various applications of computer graphics. It allows for realistic texturing of 3D objects, adding depth and detail to their appearance. Cube mapping is commonly used in video games, virtual reality, and computer-generated imagery to create immersive environments and enhance visual quality.

While cube mapping is a powerful technique, it does have some limitations. The fixed geometry of the cube may not accurately represent more complex shapes, leading to distortions or stretching in the mapped image. Additionally, the resolution of the image and the quality of the mapping can impact the visual fidelity of the result.

In conclusion, mapping an image onto a cube surface in computer graphics is a process that involves projecting a 2D image onto the six faces of a 3D cube. This technique, known as cube mapping, is widely used to create realistic textures and environment mapping in 3D rendering. It provides a versatile and effective means of adding visual details to 3D objects and environments, enhancing their realism and immersive quality in various applications.

2.1.3 Viewing Parameters

In computer graphics rendering, various visualization parameters play a crucial role in determining the appearance and quality of the rendered image. These parameters control factors such as lighting, shading, material properties, and camera settings.

The visualization parameters in computer graphics rendering include *Lighting* can be focused in a direction like the sun to provide *global illumination* or focused on spot that vary in intensity and colour will affect how objects are illuminated in the scene to provide *local* *effects*. Depending on the viewer location they may see *shadows* that add depth and realism to a rendered image. This appearance of shadows can be gradual (soft) or instant (hard) use shadow maps, or in high end ray tracing.

Lighting allows an object to be visible, but how is it seen depends on the *Shading.* This shows how objects appears depending on its *Material*, and has properties like how much it reflects colour, level of transparency, roughness that refracts light within the scene. These surface properties are calculated depending on its shading models to decide how light is scattered.

The appearance of objects is improved through *Textures* when applied to a 3D object through a process known as UV *Mapping.* A texture can be seen many times on a texture (*tiling)*, or *rotated* provides effects such as waves, or filtering out colours. These effects when combined with bump maps and face normal helps the object to reflect the direction of the face can help adjust the appearance without adjusting the model.

When viewing models in computer screens can sometimes reveal jagged lines, known as aliasing. The process of reducing these lines is *anti-aliasing*. This uses an approach such as *sampling* that uses multiple colour samples per pixel and then averaging them to create a new colour that reduces the visibility of jagged lines. Approaches for sampling such as such as super-sampling, multi-sampling, or post-processing.

Removing jagged lines is an example of *post processing* that can also be applied to create other effects such as depth of field, motion blur, post processing to stylise the images. These and other visualization parameters are adjusted by artists, designers, and engineers to achieve the desired appearance and visual fidelity in computer graphics rendering. They allow for the creation of realistic scenes, stunning visual effects, and compelling virtual environments across various applications, including gaming, animation, architectural visualization, virtual reality, and scientific simulations.

2.2 Camera and Display Systems

2.2.1 Digital Imaging and 360° Cameras

Concurrently with the development of the latest VR headsets, there has been that of the 360 cameras [7], [8]. The reason being VR headsets naturally fit with omnidirectional viewing, through head-rotation. The most interesting type of such camera systems (and also the most expensive), are now capable of acquiring stereoscopic-3D (S3D images) [9], [10], [11]. S3D is a viewing capability VR headsets naturally support through separate displays.

Latest developments in 360/3D cameras have made easier the capture of compelling photorealistic views for VR use, with great potential towards providing high-level visual realism. This has made easier the adoption of photographic and video to represent VR environments, opening up to wide use of photorealistic VR environments, which is in contrast with so far mainstream computer graphics representations. Because of this we find most literature works mainly assessing VR visual realism over synthetic images. Virtual environment navigation and interaction with VR headsets still remain challenging, while several solutions have been proposed, typically through touch controllers [12], [13], free-hand gestures [14], [15], eye fixation [16].

Parallel to latest VR headsets development, there has been that of 360 cameras. The reason being VR headsets naturally fit with omnidirectional viewing, through head-rotation. The most interesting type of such camera systems (and also the most expensive) are now capable of acquiring stereoscopic-3D (S3D) images, e.g. *Insta360 Pro*. S3D is a viewing capability VR headsets naturally support through separate displays. A 360 camera view can also be generated with standard (not 360) 2D/3D cameras, but this calls for acquisition and processing of several photographs [17], [18].

Latest developments in 360/3D cameras have made easier the capture of compelling photorealistic views for VR use, with great potential towards providing high-level visual realism. This has made easier the adoption of photographic and video to represent VR environments, opening up to wide use of photorealistic VR environments, which is in contrast with so far mainstream computer graphics representations. Because of this we find most literature works mainly assessing VR visual realism over synthetic images. Virtual environment navigation and interaction with VR headsets still remain challenging, while several solutions have been proposed, typically through touch controllers, free-hand gestures, eye fixation.

Virtual reality often benefits from panoramic camera images. This is especially the case when a sense of remote presence is sought into real environments. This aspect is relevant also within dashboard visual display as camera systems have become ubiquitous and of increasing popularity. Recently quality 360-degree camera have been spreading, which now also include high-resolution imaging, 3D and video capture. The proposed dashboard HMD based concept includes the use of remote image observation. Therefore, to familiarize with this technology has deemed relevant for my PhD work, and a few activities have been performed in relation to it.

2.2.2 Camera Parameters

Capturing the visual environment correctly is important to achieve better results within computer graphics and computer vision. If the visual environment cannot be represented correctly then decisions based off this visual information may be incorrect or not accurate and will affects perception. What this also means is that this information is unlikely to be generalisable.

A camera uses these parameters to create digital output. These parameters control the viewpoint, framing, and perspective of the rendered image. A camera has a *position* in environment and its *orientation*. A parameter to control how much of the environment is visible is via the *lens.* It has a focal length, which is the distance between the camera lens and image sensor affecting perceived distance between objects. A longer focal length results in a zoomed in effect with a narrow *field of view* and a short focal length gives a wide *field of view*. The *aperture* controls the focus, this is varying the amount of light into the camera sensor.

Additionally, the aperture also controls the *Depth of Field (DoF).* This *controls* what objects appear in sharp focus. A shallow or large aperture of 1.4 focuses on a specific close distance, blurring objects in front or behind. A small aperture of 22 is a deep setting that allows sharp focus across a broader range of distances.

Some cameras can capture stereo images that are two images taken with a horizontal distance to represent human eyes. These left and right images can then be viewed by the left and right screens in the VR headset allowing the human to then achieve binocular vision that provide depth perception and other cues.

These camera parameters enable artists, designers, and developers to control the perspective, framing, and visual storytelling of the rendered image. By adjusting these parameters, one can create realistic and visually pleasing results with different viewpoints, achieve specific visual effects, and convey the desired atmosphere or narrative in computer-generated imagery, virtual reality experiences, or visual simulations. Additionally various applications, including gaming, animation, film production, architectural visualization, and scientific simulations.

2.2.3 Display Parameters

Display parameters influence the visual quality, clarity, and overall viewing experience. One of the big factors is *Display Size* that refers to the physical dimensions of the screen, typically measured diagonally in inches. Larger displays offer more immersive experiences. The quality of the display is given by its *resolution.* The higher the resolution, the sharper and more detailed the images.

This detail is provided by *pixels,* short for picture element. These pixels can provide a certain colour range called *colour depth.* This determines the number of colours that can be displayed on the screen by a pixel. Common colour depths include 8-bit (16.7 million colours) and 10-bit (1.07 billion colours). Higher colour depth provides more accurate and vibrant colour representation.

These pixels have *Brightness* referring to light output of pixel, typically measured in nits (cd/m^2). Higher brightness levels improve visibility, especially in well-lit environments. VR Displays display differently to other displays. Normally colours are displayed in an achromatic RGB approach but can be displayed differently in VR [19]. Additionally, when pixels are displayed the brightness variation provides contrast. This *contrast ratio* results in better differentiation between dark and light areas that leads to more visually appealing images. The greater number of pixels known as within an inch known as *pixel density* provides a shaper viewing experience.

These pixels updated many times a second known as the *refresh rate* measured in Hertz (Hz). Higher refresh rates, such as 60Hz, 120Hz, or 240Hz, result in smoother motion and reduce motion blur. Lower response times reduce motion blur and ghosting in fast-paced content, such as games or videos.

This display consists of pixels can be arranged in a certain *Aspect Ratio. This is* the proportional relationship between the width and height of the rendered image. Common aspect ratios include 16:9 (widescreen) and 21:9 (ultrawide). Different aspect ratios are suitable for specific displays, content types and applications.

These display parameters affect the visual quality and user experience across various applications, including gaming, multimedia consumption, design work, and productivity tasks. When selecting a display or calibrating settings, considering these parameters helps ensure a best viewing experience tailored to specific needs and preferences.

2.3 Virtual Reality Technology and Displays

Depth perception is a fundamental aspect of human vision that allows us to perceive the world in three dimensions. It enables us to accurately gauge the distances and spatial relationships between objects, which is crucial for interacting with our environment. In the Virtual Reality Technology and Displays.

2.3.1 User Visual Interface

Virtual reality (VR) user interfaces (UIs) are crucial in enabling users to interact with and navigate within virtual environments. These interfaces play a significant role in ensuring an intuitive and immersive experience. In this two-page summary, we will explore the key aspects of VR user interfaces, including *interaction techniques, navigation methods, and design considerations.*

Interaction Techniques

Interaction techniques in VR UIs determine how users can interact with virtual objects, manipulate elements, and perform actions within the virtual environment. Various input methods and gestures are used to facilitate user interaction and engagement. Common interaction techniques in VR UIs include:

- *Controller-based Input*: VR systems often utilize handheld controllers that allow users to interact with the virtual environment. These controllers may have buttons, triggers, touchpads, or joysticks, which users can manipulate to select, grab, move, or manipulate objects.
- *Hand Tracking*: Some VR systems incorporate hand tracking technology, enabling users to interact with virtual objects using their natural hand movements. This technique enhances immersion and eliminates the need for physical controllers.
- *Gesture Recognition*: VR UIs can interpret specific hand or body gestures to trigger actions or perform tasks. Users can use gestures like pointing, grabbing, swiping, or thumbs-up to interact with virtual objects or navigate menus.

Navigation Methods

Navigation within the virtual environment is crucial for users to explore and move around virtual spaces. Effective navigation methods ensure smooth and comfortable movement, preventing motion sickness and enhancing the overall user experience. Common navigation methods in VR UIs include:

- *Teleportation*: Teleportation allows users to instantly move to a different location within the virtual environment. It reduces the risk of motion sickness associated with continuous movement and provides a comfortable and controlled way to navigate large virtual spaces.
- *Smooth Locomotion*: Smooth locomotion involves allowing users to move within the virtual environment using natural movements, such as walking or running. This method provides a more immersive experience but may require additional techniques to mitigate motion sickness, such as reducing latency and maintaining a consistent frame rate.
- *Point-and-Click*: Point-and-click navigation involves pointing at a location or object using a controller or hand tracking and selecting it to move or interact. This method is suitable for more precise navigation and interaction within the virtual environment.

Design Considerations

Designing effective VR UIs involves considering various factors to optimize usability, comfort, and immersion. The following considerations are essential for creating compelling and user-friendly VR interfaces:

- *Visual Clarity*: Text, icons, and UI elements should be designed with legibility and readability in mind. High contrast, appropriate font sizes, and clear visuals enhance readability in VR environments.
- *Depth and Layering*: Proper depth and layering of UI elements help users perceive the hierarchy of information and interact with objects at different distances. This technique provides depth cues and prevents occlusion of important UI elements.
- *Feedback and Affordances*: Providing visual and auditory feedback in response to user interactions enhances the sense of presence and confirms that actions have been successfully performed. Visual cues like highlighting or animation and audio cues can provide valuable feedback.
- *Comfortable Interactions*: VR UIs should aim to minimize user fatigue and discomfort. Considerations such as button placement, interaction techniques, and the need for repetitive actions should be taken into account to ensure comfortable and ergonomic interactions.

VR UIs continue to evolve with advancements in technology and user interface research. Some promising developments include:

- *Natural Language Processing*: Integration of voice recognition and natural language processing capabilities can enable users to interact with virtual environments using spoken commands, enhancing convenience and immersion.
- *Artificial Intelligence and Machine Learning*: AI and ML algorithms can improve VR UIs by understanding user behavior, preferences, and intent. Personalized UIs that adapt to individual users' needs and provide tailored experiences are potential applications.

2.3.2 Depth Cues and Stereoscopic 3D Visualization

Depth perception is a fundamental aspect of human vision that allows us to perceive the world in three dimensions. It enables us to accurately gauge the distances and spatial relationships between objects, which is crucial for interacting with our environment. In the context of visual media, such as movies, gaming, and virtual reality (VR), creating a sense of depth is essential for a realistic and immersive experience. Depth cues and stereoscopic 3D viewing techniques are employed to simulate depth perception and enhance the perception of depth in these visual mediums.

We see the world in 3D generated by two 2D images collected from the back of our eyes. The depth information created uses a process in the mind called stereopsis in the visual part of the brain. Stereopsis is the resulting sense of depth when two slightly different views fused into one 3-dimensional image conceptually shown in the figure below.

Figure 5 - The advantage of 3D vision over 2D vision

To understand stereopsis and artificial 3D vision means learning about the brain visual pathways (Figure 6) and how human visual system works. It uses cues from the images from the eye that induce perception of depth by means of monocular and binocular cues.

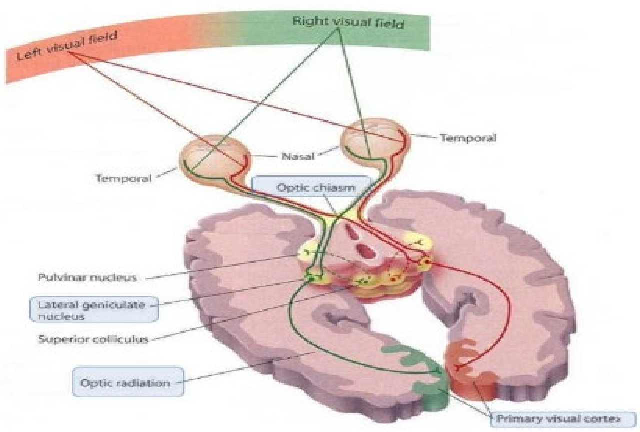


Figure 6 - Brain Visual Pathways

Depth Cues are visual cues or hints that our brain uses to interpret depth and distance in a scene. These cues help us perceive objects as being closer or farther away from us. There are monocular cues and binocular cues.

Monocular cues convey some depth information even in standard two-dimensional images, and consists of *Position, Perspective, Relative Size, Aerial Perspective, Defocus Blur, Light and Shade, Texture Gradients, Occlusions, Motion Parallax*. Figure 7 shows some monocular cues.

Figure 7 - Monocular cues.

Monocular cues rely on visual information that can be perceived with one eye alone include *Perspective* that indicates that it is farther away and appears smaller as converges towards a vanishing point on the horizon. If the objects have a *shadow* gives an idea of distance and position. When this is combined with *texture gradient* it can show that the object is in the distance when the texture is has less detail. If object *overlaps* with another it is possible to understand what object is in front. As an object moves in the distance it will appear to move slow, and if close it will appear to move fast that is called *motion parallax*.

Binocular cues rely on *binocular convergence* and *binocular parallax*. When the eyes focus on an object, they can change rotation angle to guarantee a clear visualization of the gazed point. When the focus is on the wrong point this cause Defocus blur. This eye angle rotation defines convergence used by the visual system to judge distance. The angle of convergence is smaller when the eye is fixating on far away objects.

In real life, convergence combined with accommodation improves distance estimation and focus on the subject. On a 2D display these are in conflict. The eyes converge on the observed object, but accommodation does not change on a 2D screen.

Binocular parallax is the viewing of object viewed slightly differently between left and right views creating disparity that induces stereopsis and simulates depth perception (Winkler & DongboMin). This also supplies a better perspective view. Figure 8 left image shows eye focus varies based on object distance and the right image shows different in left and right image.

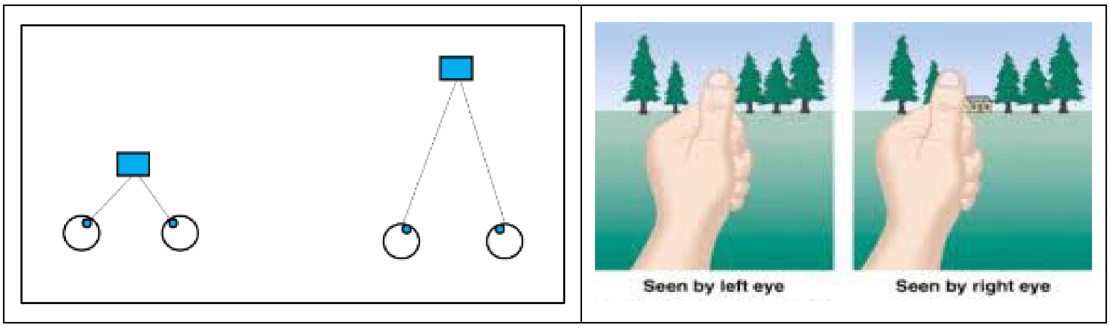


Figure 8 - Scheme of Binocular convergence (Left) and Binocular Parallax (Right) - Binocular cues.

Display technologies provide S3D and related cues to various quality has been in development over last the two centuries with few systems developed. Within the last thirty years there are three popular S3D display technologies:

Passive stereo, which makes use of space-multiplexed images that can be visualized using *anaglyph glasses using with colour filters, polarized glasses with polarized filters, or separated displays* located very close to viewer's eyes (e.g. HMDs) shown in Figure 9.

Figure 9 - Anaglyph glasses. Polarized glass. Separated displays.

Active stereo, which uses time-multiplexed images can be visualized through *shutter glasses* offering liquid crystal display (LCD) shutter panels synchronized with the visualization display as shown in Figure 10.

Figure 10 - Nvidia Shutter glasses. Sony active 3D glasses.

Autostereoscopic stereo, uses special reflecting layers with *parallax barriers* and *lenticular sheets* to provide the correct 3D perspective of the virtual environment depending on viewing position avoiding the need for googles as shown in Figure 11

Figure 11 - Lenticular Sheets.

2.3.3 Virtual Reality Displays

Virtual reality (VR) displays are advanced technologies that immerse users in realistic and interactive virtual environments. These displays utilize a combination of hardware and software to create a sense of presence, allowing users to feel as if they are physically present within the virtual world. Here is a summary of VR displays:

Display Technologies

- *Head-Mounted Displays (HMDs)*: These are the most common type of VR displays. HMDs consist of a head-worn unit that houses one or two screens, positioned in front of the user's eyes. They often incorporate additional features such as built-in sensors for tracking head movements, built-in audio systems, and sometimes even haptic feedback.
- *Cave Automatic Virtual Environments (CAVEs)*: CAVE systems consist of large-scale, room-sized displays where users are surrounded by screens or projectors on multiple walls. The immersive experience is achieved by projecting the virtual environment onto these walls.
- *Augmented Reality (AR) Displays*: AR displays overlay virtual objects onto the real-world environment, allowing users to interact with both the physical and virtual elements simultaneously. AR displays can be HMDs or transparent screens that overlay virtual content onto the user's field of view.

Tracking and Input Devices

- *Positional Tracking*: To enable realistic movement within the virtual environment, VR displays employ various tracking systems. These systems track the user's head movements, allowing the display to adjust the perspective accordingly. Common tracking methods include infrared sensors, gyroscopes, accelerometers, and external cameras.

- *Controllers*: VR displays often come with handheld controllers that allow users to interact with the virtual world. These controllers can track hand movements, gestures, and button inputs, providing a more immersive and intuitive way to manipulate objects and navigate the virtual environment.

Display Resolution and Refresh Rate

- *High Resolution*: VR displays require high-resolution screens to provide clear and sharp visuals. Higher resolution reduces the screen-door effect, where users see a grid-like pattern due to pixelation. Common resolution standards include Full HD (1920x1080 pixels) or higher, and some displays offer even higher resolutions for enhanced visual fidelity.
- *Refresh Rate*: VR displays demand high refresh rates to minimize motion sickness and provide smooth, responsive experiences. Typical refresh rates range from 90Hz to 120Hz, ensuring that the display updates quickly enough to match the user's head movements, reducing lag and motion blur.

Optics and Field of View

- *Optics*: VR displays employ various lens systems to focus the virtual images onto the user's eyes. These lenses help increase the perceived depth and immersion by enlarging the virtual scene and enhancing the visual quality.
- *Field of View (FoV)*: FoV refers to the extent of the virtual environment visible to the user. A wide FoV provides a more immersive experience, as it fills the user's peripheral vision. VR displays aim to achieve FoVs similar to or greater than the human vision (around 100 to 120 degrees).

Software and Content

- *VR Software Development Kits (SDKs)*: SDKs provide tools, libraries, and APIs that enable developers to create VR applications and experiences. Popular SDKs include Unity, Unreal Engine, and SteamVR.
- *Content and Applications*: A thriving ecosystem of VR applications has emerged, including gaming, educational experiences, simulations, training programs, and virtual tours. Both entertainment and practical applications benefit from the immersive nature of VR displays.

Virtual reality displays offer a transformative and engaging experience by immersing users in virtual worlds. With the continual advancements in technology, VR displays are becoming more accessible, affordable, and capable of delivering realistic and compelling experiences across various industries and domains.

Different display systems can be used to visualize S3D images. These are responsible for the degree of immersion, isolation from the surrounding environment, and image quality of the observed photographs. One of the earliest immersive virtual reality displays is the *stereoscope*, which was devised by Wheatstone in 1838 [20] [21]. This device uses mirrors (see Figure 12) to show on each eye a different perspective of the observed scene, and therefore induce stereopsis by binocular parallax providing depth perception.

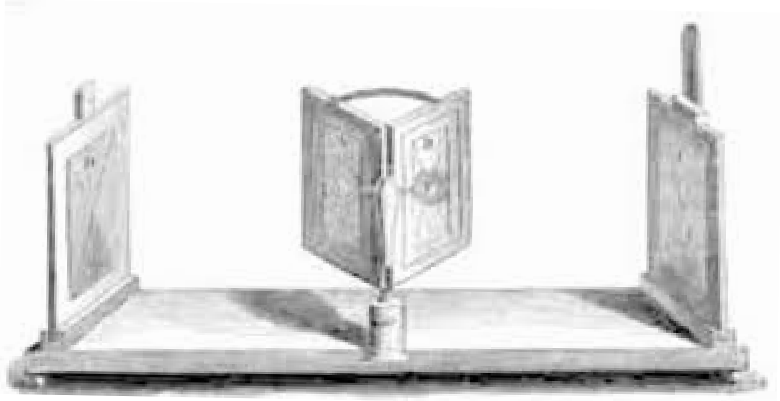


Figure 12 - Concept representing Wheatstone's stereoscope.

We can distinguish display systems show in Figure 13:

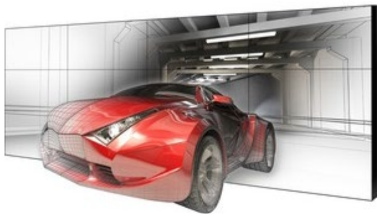
Desktop / Laptops. Which are suitable for the active and passive stereo 3D approaches, and offer CRT or LCD displays	
Walls. Which offer large screens with different projection modalities and different display structures	
Rooms. Which can provide cubic or multi-part screens to enable high user involvement (e.g. 3D CAVEs)	
Handheld. Which are portable and are suitable for the passive stereo 3D approach (e.g. Smartphones, Tablets, LED screens)	
Head Mounted Displays (HMDs). Which offer very small displays located inside portable helmets at very short distance from the viewer, enabling higher levels of isolation from the surrounding space (e.g. HMDs).	

Figure 13 - 3D Display Systems

2.3.4 VR Headsets – HMDs

Ivan Sutherland built the first HMD prototype in 1968. VR headsets have since then received renown interest because of their potential in providing full visual immersion into artificially generated environments. However, up until last decade, HMDs use had mainly be confined within research labs. The main reasons being: cost, weight, portability and display performance.

While relevant improvements occurred in the last decades, with HMDs adopting optical tracking and OLED displays, while becoming smaller and lighter, some other issued remained, e.g. tunnel vision, PC tethering and high cost. These issues had limited HMD adoption into the consumer market. A notable step forward was the first *Oculus Rift* system in 2012, featuring wide FOV and low-cost.

There has been great development since then, with newer systems featuring wider displays and higher resolution, lower cost and wireless connection. The latter has been a major focus on latest systems. Smartphone-based VR headsets have first been proposed, which rely on smartphone's hardware and its display to operate, e.g. *Google Cardboard* and *Samsung GearVR*. Hence, standalone VR headsets have come into the market, which rely on dedicated computing hardware and display, e.g. *Oculus Quest*.

The highest-specs VR headset have nonetheless remained those wired to desktop PCs, because they can exploit greater processing and graphic power, e.g. *Oculus Rift* and *HTC Vive*. Recent enhancements to VR headsets include embedded head and eye trackers, e.g. *HTC Vive Pro Eye*. The figure below shows examples of different types of VR headsets, including wireless, wired and smartphone based. The latter allows us to assess performance with different displays.

Figure 14 - Examples of VR headsets (wireless, wired and smartphone based) allowing for omnidirectional stereoscopic-3D viewing. The image top-left shows the Oculus Quest 2 and the top-right the HTC Vive Pro. The bottom row shows smartphone-based headsets (Samsung GearVR and Google DayDream).

The interest towards HMDs has recently got new momentum because of the technology advances in terms of high field-of-view display. HMDs (also referred as VR Headsets) well meet the needs that virtual reality observation have, to enhance quality of remote observations. HMDs can be classified into:

Desktop VR headsets

Desktop VR headsets require high computations and a wired connection to an auxiliary computer or laptop that offers high-end graphic cards such as NVIDIA GeForce RTX 3080 Ti [22], (see Figure 15).

Figure 15 - HMDs that require PC- HTC Vive, Oculus Rift Fove, Play Station VR, StarVR and PIMAX .

Mobile VR headsets

Mobile VR headsets require a smartphone to be used as a display to show stereoscopic content, and are portable with no wires (see Figure 16).

Advantages:

- High Portability
- No wires needed
- Easy to mount
- Cheap
- Works with smartphones

Disadvantages:

- Performances limited by smartphones
- Sometimes poor quality lenses
- Frequent delays with rich content

Figure 16 - HMDs which require a smartphone - Google Cardboard and Gear VR.

Stand-Alone VR headsets

Stand-Alone VR headsets have no needs for further devices to work and can offer an integrated operative system to interact within virtual reality, (see Figure 17).

Advantages:

- High Portability
- Good Performances
- Good quality lenses
- No wires needed
- Easy to mount
- Cheap

Disadvantages:

- Frequent delays with rich content
- Low Memory storage

Figure 17 - Stand-Alone HMDs - Oculus Quest 2 and HTC VIVE Focus.

Latest VR Headset Production

The virtual reality (VR) headset market is continually evolving, with several prominent brands and models offering a range of features and experiences. Here is a summary of some of the latest VR headset models and brands:

- *Oculus Quest 2*: The Oculus Quest 2, developed by Facebook's Oculus, is a popular standalone VR headset. It offers high-resolution displays, inside-out tracking, and a wide library of VR games and applications. With its wireless capabilities, it provides an untethered VR experience and offers different storage options to suit varying needs.

- *HTC Vive Pro 2*: The HTC Vive Pro 2 is a premium PC-powered VR headset known for its high-resolution displays, delivering sharp visuals and immersive experiences. It offers precise tracking, adjustable head strap, and compatibility with the SteamVR ecosystem, making it popular among gamers and VR enthusiasts.
- *Valve Index*: Developed by Valve Corporation, the Valve Index is a high-end VR headset designed for PC gaming. It features industry-leading tracking technology, high refresh rates, and a wide field of view. The Valve Index offers precise hand-tracking with its controllers, making it suitable for immersive gaming experiences.
- *PlayStation VR*: Developed by Sony Interactive Entertainment, PlayStation VR is a popular VR headset designed for use with PlayStation 4 and PlayStation 5 gaming consoles. It offers a wide range of games and experiences, comfortable ergonomics, and easy setup. PlayStation VR is a cost-effective option for console gamers looking to enter the world of VR.
- *HP Reverb G2*: The HP Reverb G2 is a high-resolution PC-powered VR headset developed in collaboration with Valve and Microsoft. It boasts impressive visual quality, with sharp displays and enhanced clarity. The headset offers inside-out tracking, comfortable design, and compatibility with a wide range of VR applications and games.
- *Pico Neo 3*: The Pico Neo 3 is a standalone VR headset with a focus on enterprise and business applications. It offers a lightweight design, 6DoF tracking, and powerful hardware specifications. The Pico Neo 3 is often used for training simulations, virtual tours, and other professional use cases.
- *Samsung Odyssey Series*: The Samsung Odyssey series includes PC-powered VR headsets, such as the Odyssey+ and the Odyssey G7. These headsets feature high-resolution displays, comfortable designs, and compatibility with Windows Mixed Reality. They provide an immersive VR experience for gaming, entertainment, and productivity applications.

These are just a few examples of the latest VR headset models and brands available in the market. Each headset offers unique features, specifications, and target audiences, catering to different needs and preferences. The VR headset market continues to expand, with ongoing advancements in technology and an increasing focus on improving comfort, resolution, tracking, and overall immersion.

2.3.5 Virtual Reality Software

Virtual reality (VR) software refers to computer programs and applications that enable users to immerse themselves in a simulated digital environment. It utilizes specialized hardware, such as VR headsets, to create an interactive and three-dimensional experience that can be explored and interacted with by users in a more immersive and realistic way than traditional computer interfaces.

VR software typically consists of several components including: VR Rendering, Tracking and Input, Interaction and User Interface, Audio and Sound Effects, Content and Applications, Integration and Development Tools.

There several software in the market proposed to develop virtual reality applications and to describe virtual reality environments. These are typically represented by graphics software but they also integrate audio and other sensory inputs such as haptics.

Game engines such as *Unity3D* (Unity_Technologies, 2019) and *Unreal* (Unreal_Engine, s.d.). represent recent commercial products that have largely been used to design virtual reality environments and human interaction. Unity and Unreal Engine are two leading virtual reality (VR) software platforms that empower developers to create immersive and interactive VR experiences. Here's a one-page summary highlighting the key features and strengths of each platform:

Unity is a widely used VR software platform known for its user-friendly interface and versatility. It offers a comprehensive suite of tools and features that make VR development accessible to developers of all levels of expertise.

It offers Cross-platform Support: Unity supports multiple VR platforms, including Oculus Rift, HTC Vive, Windows Mixed Reality, PlayStation VR, and more. This cross-platform compatibility enables developers to reach a broad audience and ensures flexibility in targeting different VR hardware.

Real-time Rendering: Unity's powerful real-time rendering engine allows for the creation of visually stunning and interactive VR environments. It supports advanced graphics features, including dynamic lighting, post-processing effects, and high-fidelity shaders, contributing to a realistic and immersive VR experience.

Unity provides robust tools for implementing VR interaction and input systems. Developers can utilize Unity's built-in support for hand tracking, motion controllers, and spatial mapping to enable users to interact with virtual objects and navigate the VR environment, enhancing immersion and user engagement.

Unreal Engine, is renowned for its powerful graphics capabilities and is widely used in the creation of high-fidelity and visually stunning VR experiences. Unreal Engine excels in generating realistic and visually impressive VR environments. Its advanced rendering techniques, including dynamic lighting, global illumination, and physically based materials, contribute to the creation of immersive VR visuals that closely resemble real-world scenarios.

Unreal Engine's Blueprint visual scripting system simplifies the creation of interactive VR experiences without extensive programming knowledge. Developers can visually connect nodes to create VR interactions, gameplay mechanics, and complex behaviours, enabling designers and artists to actively participate in the VR development process.

Tools: Unreal Engine provides a range of VR-specific development tools to facilitate VR application creation. These tools include a VR Editor that allows developers to build and modify VR environments directly within VR, as well as workflows for tasks like level design, optimization, and performance profiling.

Unreal Engine offers built-in support for various VR motion controllers, allowing developers to easily integrate and map controller input to in-game interactions. It also supports different tracking systems, ensuring accurate tracking of user movements within the VR space for a seamless and responsive experience.

Both Unity and Unreal Engine have made significant contributions to the advancement of virtual reality software. They provide developers with the necessary tools, capabilities, and resources to create immersive and engaging VR applications across various platforms. The choice between Unity and Unreal Engine often depends on factors such as personal preference, project requirements, and the specific features and strengths offered by each platform.

2.4 Immersion and Human Factors

2.4.1 The Immersion Concept

Immersion is a fundamental concept in virtual reality (VR) that refers to the degree of realism and engagement experienced by a user within a virtual environment. It aims to create a sense of presence, where the user feels fully absorbed and connected to the virtual world, disconnecting from the physical surroundings. Achieving a high level of immersion is essential in providing a compelling and convincing VR experience. Here is a summary of the concept of immersion in VR:

Sensory Immersion

- *Visual Immersion*: Visual immersion involves high-quality and realistic graphics, including accurate rendering, lighting, textures, and detailed environments. The visuals should provide a convincing representation of the virtual world, minimizing visual artifacts and discrepancies.
- *Auditory Immersion*: Sound plays a crucial role in creating an immersive experience. Realistic and spatial audio enhances the sense of presence by accurately reproducing sounds based on the user's position and environment within the virtual world.
- *Haptic Immersion*: Haptic feedback, such as vibrations or tactile sensations, enhances immersion by providing users with a sense of touch or physical interaction within the virtual environment. Haptic devices can simulate textures, forces, and interactions, further immersing the user in the experience.

Interaction and Control

- *User Interaction*: Immersion in VR is enhanced by providing users with intuitive and responsive methods of interaction. This includes hand-tracking, gesture recognition, motion controllers, and other input devices that allow users to manipulate objects and interact with the virtual environment.
- *Control and Navigation*: Seamless and natural control mechanisms, such as head tracking and body movement tracking, contribute to the feeling of immersion. These

methods enable users to explore and navigate the virtual space in a way that mimics real-world movements.

Real-Time Responsiveness

- *Low Latency*: Achieving low latency between user actions and corresponding system responses is vital for maintaining immersion. Delayed responses can break the sense of presence and disrupt the user's engagement. High-performance hardware and efficient software algorithms are essential for minimizing latency in VR systems.

The concept of immersion is central to virtual reality, aiming to transport users to believable and engaging virtual worlds. By integrating realistic visuals, accurate audio, responsive interactions, and a sense of presence, VR experiences can captivate users' senses and emotions, blurring the line between the physical and virtual realms. Advancements in technology, including graphics rendering, tracking systems, haptic devices, and user interfaces, continue to push the boundaries of immersion in VR, enabling increasingly realistic and immersive experiences for a wide range of applications, including gaming, training simulations, education, and virtual tourism.

2.4.2 Sense of Presence, Realism, Comfort and Emotions

Human factors play a significant role in the design and development of virtual reality (VR) systems, particularly when it comes to aspects such as presence, realism, comfort, and emotions. These factors greatly influence the overall user experience and immersion within the virtual environment. In this four-page summary, we will delve into each of these areas and explore how human factors contribute to their optimization in VR.

Presence

Presence refers to the feeling of "being there" in the virtual environment, where users experience a sense of believability and immersion. Achieving a strong sense of presence is crucial for creating an engaging and realistic VR experience. Human factors considerations for presence in VR include:

- *Visual Realism*: High-quality graphics, accurate rendering, and realistic lighting are essential to create visually convincing virtual worlds. Detailed textures, accurate object scale, and appropriate levels of detail contribute to a more immersive experience.
- *Audio Immersion*: Realistic and spatial audio cues enhance the sense of presence by accurately replicating sounds from different directions. Proper sound design, including 3D audio techniques and accurate sound propagation, contribute to a more immersive and realistic auditory experience.
- *Interaction and Feedback*: Intuitive and responsive interactions with virtual objects, environments, and characters help create a sense of presence. Real-time haptic feedback and precise motion tracking improve the user's sense of being physically connected to the virtual world.

Realism

Realism in VR refers to the extent to which the virtual environment resembles the real world. The more realistic the experience, the more immersed and engaged users become. Human factors considerations for realism in VR include:

- *Visual Fidelity*: High-resolution displays, accurate color reproduction, and realistic lighting techniques enhance visual realism. Proper attention to shading, reflections, and environmental effects can significantly contribute to a more convincing and immersive virtual environment.
- *Physics and Interactions*: Realistic physics simulations and accurate object interactions help users suspend disbelief and feel connected to the virtual world. Objects should behave naturally, responding to forces and collisions in a way that aligns with real-world expectations.
- *Environmental Cues*: Environmental cues, such as weather effects, atmospheric conditions, and spatial soundscapes, contribute to a more realistic and immersive experience. Paying attention to these details helps users establish a stronger connection to the virtual environment.

Comfort

Comfort is a critical factor in VR experiences, ensuring that users can engage in VR activities for extended periods without discomfort or adverse effects. Various human factors considerations contribute to user comfort in VR. Human factors considerations for comfort in VR include:

- *Ergonomics*: VR hardware, such as headsets and controllers, should be designed with ergonomic considerations in mind. Proper weight distribution, adjustable straps, and padding help minimize discomfort during prolonged use.
- *Motion Sickness Mitigation*: Motion sickness can occur due to discrepancies between perceived motion and actual motion. Reducing latency, optimizing frame rates, and employing smooth locomotion techniques, such as teleportation or incremental movement, can help minimize motion sickness and enhance user comfort.
- *Optimal Field of View*: Providing users with an appropriate field of view helps reduce visual discomfort and fatigue. A wide field of view can enhance immersion, while narrow fields of view may cause tunnel vision or disorientation.

Emotions

Emotional engagement is an important aspect of VR experiences, as it helps users establish a deeper connection to the virtual world and elicit specific emotional responses. Human factors considerations contribute to the emotional impact of VR. Human factors considerations for emotions in VR include:

- *Storytelling and Narrative*: Crafting compelling storylines, characters, and scenarios can evoke emotional responses from users. Engaging narratives, well-developed characters, and dynamic events or challenges contribute to emotional engagement and enhance the overall experience.

- *Music and Sound Design*: Music and sound effects play a significant role in shaping emotions.

2.4.3 Depth Perception and Distance Estimation

Human factors in virtual reality (VR) play a critical role in creating a convincing and immersive experience, particularly when it comes to depth perception and distance estimation. These factors contribute to the user's ability to perceive and interact with the virtual environment accurately. In this four-page summary, we will explore the importance of human factors in relation to depth perception and distance estimation in VR.

Depth Perception in VR

Depth perception refers to the ability to perceive the spatial relationships between objects in a three-dimensional space. It enables users to judge the distance, size, and position of objects accurately. In VR, depth perception is crucial for creating a sense of realism and spatial understanding. Human factors considerations for depth perception in VR include:

- *Stereoscopic Vision*: VR systems often utilize stereoscopic displays, which present slightly different images to each eye to create a perception of depth. By mimicking the natural binocular vision of humans, stereoscopy enhances depth perception in VR.
- *Convergence and Accommodation*: Convergence is the inward movement of the eyes to focus on a specific point, while accommodation is the adjustment of the eye's lens to focus on objects at different distances. In VR, the use of proper convergence and accommodation cues helps simulate natural depth perception.
- *Size and Distance Cues*: Virtual objects should be appropriately scaled and positioned to provide accurate size and distance cues. Familiar size cues, such as objects becoming smaller with distance, can assist users in perceiving depth accurately.

Distance Estimation in VR

Accurate distance estimation is essential in VR to create a realistic sense of scale, object placement, and spatial relationships. Users' ability to perceive distances correctly enhances the overall immersion and interaction within the virtual environment. Human factors considerations for distance estimation in VR include:

- *Motion Parallax*: Motion parallax refers to the relative movement of objects as the user moves their head or body. Incorporating motion parallax in VR provides important depth and distance cues, as objects closer to the user will appear to move faster than objects in the distance.
- *Texture Gradient*: The texture gradient describes how the size and density of a texture change as it recedes into the distance. Utilizing texture gradient cues in VR environments assists users in estimating distances accurately.

- *Shadows and Lighting*: Proper use of lighting and shadows in VR can help users perceive depth and distance. Shadows cast by objects can provide important visual cues that aid in estimating the spatial relationships between objects.

Human Factors Challenges and Considerations

While VR systems aim to simulate depth perception and distance estimation, there are challenges and considerations to be aware of to ensure an optimal user experience.

- *Visual Discrepancies*: Discrepancies between the virtual and physical world, such as limited field of view or resolution, can affect depth perception and distance estimation. Designing VR experiences with appropriate visual cues and minimizing discrepancies can help overcome these challenges.
- *Motion Sickness*: Inaccurate depth perception or distance estimation can contribute to motion sickness in VR. It is crucial to optimize these factors to reduce the risk of discomfort and maintain user comfort during VR experiences.
- *Individual Differences*: Human factors can vary among individuals. Factors such as age, visual acuity, and previous experience with VR may affect depth perception and distance estimation. Considering individual differences and providing customization options can enhance the user experience for a broader audience.

Future Directions and Advancements

Ongoing research and technological advancements continue to improve depth perception and distance estimation in VR. Here are some notable areas of progress:

- *Eye Tracking*: Eye-tracking technology allows for more accurate depth perception and distance estimation by precisely monitoring the user's gaze. This can help optimize rendering, adjust focus depth, and provide realistic accommodation cues in VR.
- *Foveated Rendering*: Foveated rendering focuses computational resources on the user's foveal region (the central area of highest visual acuity), reducing the processing load and enhancing visual fidelity. This technology can contribute to more accurate depth perception and distance estimation in VR.
- *Haptic Feedback*: Incorporating haptic feedback, such as touch and force feedback, can enhance depth perception and distance estimation by providing additional sensory cues. Haptic devices that simulate object weight, resistance, or texture can improve users' ability to interact with the virtual environment and estimate distances more accurately.

By understanding and incorporating human factors considerations related to depth perception and distance estimation in VR, developers can create more realistic and immersive experiences. Advancements in technology, coupled with ongoing research, will continue to refine these factors, making VR environments more convincing, engaging, and comfortable for users across various applications, including gaming, education, training, and simulations.

2.5 Visual Experience Evaluation

2.5.1 Methods for Testing VR Visual Experience

Ensuring a high-quality and immersive visual experience is crucial in Virtual Reality (VR) applications. To achieve this, rigorous testing methods and user studies are employed to evaluate and optimize the VR visual experience.

Testing VR visual experience involves a combination of methods such as visual inspection, performance testing, comparative testing, and usability testing. These methods allow developers to assess visual fidelity, performance, and usability aspects of the VR environment. Additionally, user studies, including surveys, interviews, eye tracking, and physiological measurements, provide insights into user perception, satisfaction, and emotional responses to the VR visual experience. By integrating these methods, developers can ensure a visually captivating, immersive, and user-friendly VR visual experience that meets user expectations and enhances user satisfaction.

Methods for Testing VR Visual Experience

- *Visual Inspection*: Visual inspection involves a detailed evaluation of the VR environment by experienced testers. They assess factors such as image quality, resolution, color accuracy, lighting, textures, and the presence of visual artifacts. Visual inspection provides valuable insights into the visual fidelity and immersion of the VR experience.
- *Performance Testing*: Performance testing focuses on evaluating the technical aspects of the VR system. Testers measure frame rates, latency, and response times to ensure a smooth and responsive visual experience. Performance testing helps identify any performance bottlenecks that may impact the visual quality and user experience.
- *Comparative Testing*: Comparative testing involves comparing different configurations or variations of the VR visual experience. Testers may adjust parameters such as rendering techniques, texture quality, lighting conditions, or post-processing effects to assess their impact on visual fidelity and immersion. By conducting comparative testing, developers can identify the most visually appealing and effective options.
- *Usability Testing*: Usability testing assesses the effectiveness and efficiency of the VR visual interface. Testers evaluate how well users navigate menus, interact with objects, and perform tasks within the VR environment. Usability testing helps identify visual design flaws and ensures that the visual interface is intuitive, user-friendly, and visually appealing.

Importance of User Studies

User studies play a vital role in understanding user perception, satisfaction, and comfort in the VR visual experience. These studies involve gathering feedback from a group of users who interact with the VR application. Here are some common user study methodologies:

- *User Surveys*: User surveys collect subjective feedback from users regarding their visual experience, comfort levels, presence, and overall satisfaction. Surveys help identify areas for improvement, gauge user preferences, and gather quantitative and qualitative data for analysis.

- *Interviews and Focus Groups*: Interviews and focus groups allow for in-depth discussions with users, providing valuable insights into their experiences, preferences, and suggestions for enhancing the visual experience. These qualitative methods help researchers gain a deeper understanding of user perceptions and emotions.
- *Eye Tracking*: Eye tracking technology monitors and records the eye movements of users while they interact with the VR environment. This data helps analyze visual attention, gaze patterns, and areas of interest within the virtual scene. Eye tracking can provide valuable information on how users perceive and interact with visual elements.
- *Physiological Measurements*: Physiological measurements, such as heart rate, skin conductance, and electroencephalography (EEG), can provide objective insights into users' emotional and cognitive responses to the VR visual experience. These measurements help understand the impact of visuals on users' arousal, engagement, and emotional states.

User studies complement traditional testing methods by capturing the subjective experiences and preferences of users. They provide valuable feedback to developers, enabling them to refine the visual design, optimize user comfort, and enhance the overall VR visual experience.

2.5.2 Subjective Ratings

Subjective ratings play a crucial role in evaluating the Virtual Reality (VR) visual experience as they capture users' perceptions, preferences, and satisfaction levels. These ratings provide valuable insights into the effectiveness of visual elements and help guide improvements in VR design.

Subjective ratings provide valuable information about user perception, satisfaction, and feedback regarding the VR visual experience. By incorporating rating scales, surveys, interviews, and user testing, developers can gather rich data on users' subjective experiences. These ratings offer insights into visual clarity, realism, user satisfaction, and areas for improvement. By leveraging subjective ratings, developers can refine the VR visual design, enhance user engagement, and create immersive experiences that meet user expectations and preferences.

Importance of Subjective Ratings

- *User Perception*: Subjective ratings allow users to express their perception of the VR visual experience. Users can provide feedback on aspects such as visual clarity, realism, depth perception, color accuracy, and overall immersion. These ratings reflect the users' subjective experience and provide insights into how well the VR visuals align with their expectations.
- *User Satisfaction*: Subjective ratings help gauge user satisfaction with the VR visual experience. Users can rate their overall satisfaction level, comfort, presence, and

engagement. By capturing users' satisfaction ratings, developers can assess the success of the VR visuals in creating a compelling and enjoyable experience.

- *Feedback for Improvement*: Subjective ratings offer an opportunity for users to provide feedback and suggestions for improvement. Users can identify visual issues, highlight areas that require refinement, and suggest changes that could enhance the overall visual experience. This feedback is invaluable for developers to iterate and optimize the VR visual design.

Methods for Collecting Subjective Ratings

- *Rating Scales*: Rating scales, such as Likert scales, allow users to rate specific aspects of the VR visual experience on a predefined scale. Users can assign scores or indicate their level of agreement or satisfaction for each item. These scales provide quantifiable data that can be analyzed to identify trends and patterns in user ratings.
- *Surveys and Questionnaires*: Surveys and questionnaires collect users' subjective ratings through a series of structured or open-ended questions. Users can rate various aspects of the VR visual experience and provide comments or additional feedback. Surveys provide a comprehensive overview of user perceptions and opinions.
- *Interviews and Focus Groups*: Interviews and focus groups facilitate in-depth discussions with users about their subjective experiences with the VR visuals. Testers can ask probing questions and encourage participants to elaborate on their ratings, preferences, and suggestions. These qualitative methods provide rich insights into users' thoughts and emotions.
- *User Experience (UX) Testing*: UX testing involves observing users' interactions with the VR visual experience and capturing their real-time feedback. Testers can gather users' subjective ratings during or immediately after the VR session, allowing for immediate impressions and reactions to the visuals.

Analysis and Interpretation

Analysing subjective ratings involves aggregating and interpreting the collected data. Testers can identify common trends, patterns, and outliers in the ratings to derive meaningful insights. Statistical analysis, data visualization, and qualitative analysis techniques can aid in understanding the strengths and weaknesses of the VR visual experience, guiding improvements and optimizations.

2.5.3 Objective Measurements

Objective measurements are critical in assessing the Virtual Reality (VR) visual experience as they provide quantitative data and objective insights into the performance and effectiveness of visual elements. These measurements complement subjective ratings and help developers evaluate visual fidelity, performance, and user comfort. In this one-page summary, we will explore the significance of objective measurements and their use in testing VR visual experience with users.

Objective measurements play a vital role in testing the VR visual experience, providing quantitative data and objective insights. By measuring performance metrics, visual fidelity, depth perception, and visual artifacts, developers can assess the technical performance and effectiveness of the visuals. Eye tracking, biometric sensors, performance monitoring, and quality assessment tools enable the collection of objective measurements. By leveraging these measurements, developers can refine the VR visuals, optimize performance, and create immersive experiences that deliver high-quality, visually compelling, and comfortable VR visual experiences for users.

Importance of Objective Measurements

- *Performance Metrics*: Objective measurements capture technical performance aspects of the VR visual experience. These metrics include frame rates, latency, response times, and rendering times. By quantifying these parameters, developers can assess the smoothness and responsiveness of the visuals, ensuring a seamless and immersive experience for users.
- *Visual Fidelity*: Objective measurements provide insights into the visual quality and fidelity of the VR experience. Parameters such as resolution, pixel density, color accuracy, brightness, contrast, and image sharpness can be objectively measured. This data helps assess how well the VR visuals match the intended design and deliver a high-quality visual experience.
- *Depth Perception*: Objective measurements aid in evaluating the effectiveness of depth cues in VR. Parameters like depth perception accuracy, stereoscopic alignment, and interocular distance can be measured objectively to ensure accurate depth representation. These measurements are crucial in creating a realistic and immersive visual environment.
- *Visual Artifacts*: Objective measurements can detect and quantify visual artifacts, such as aliasing, motion blur, or flickering. These artifacts can impact the quality of the VR visuals and the overall user experience. Objective measurements help identify the presence and severity of visual artifacts, enabling developers to address them effectively.

Methods for Objective Measurements

- *Eye Tracking*: Eye tracking technology allows for the objective measurement of users' eye movements and gaze patterns within the VR environment. This data provides insights into visual attention, fixation durations, and areas of interest. Eye tracking can help determine the effectiveness of visual elements and assess users' engagement with the VR visuals.
- *Biometric Sensors*: Biometric sensors, such as heart rate monitors or galvanic skin response sensors, can objectively measure physiological responses to the VR visual experience. These measurements offer insights into users' emotional arousal, stress levels, and engagement with the visuals. Biometric data helps assess the impact of visuals on users' physiological responses.

- *Performance Monitoring*: Objective measurements can be obtained through performance monitoring tools integrated into the VR system. These tools capture technical metrics, such as frame rates, latency, and response times, providing real-time data on the performance of the VR visuals.
- *Quality Assessment Tools*: Specialized software tools exist for objectively assessing visual quality and detecting artifacts. These tools analyze the VR visuals, measure parameters such as resolution, pixel density, and color accuracy, and provide objective assessments of the visual fidelity.

Analysis and Interpretation

Objective measurements require careful analysis and interpretation to derive meaningful insights. Testers can compare measurements against established benchmarks or industry standards to evaluate the performance and visual quality of the VR experience. Statistical analysis, data visualization, and comparison of objective measurements with subjective ratings can help identify correlations and trends, guiding improvements and optimizations.

3. Impact of Imaging and Distance Perception in VR Immersive Visual Experience

This PhD project aimed at assessing and understanding the VR immersive experience particularly focusing on the visual output. The first was proposed to better understand the key factors a VR visual experience relies on. The study follows up based on the lesson learnt by proposing and assessing the photo-based VR visual experience.

Because of the complexity of assessing the VR immersive visual experience, there was a need to delimit the research study. The focus was then set to the impact of imaging and distance perception.

In this chapter the PhD project core idea and motivation are first presented. The three main phases of study, development, and experimentation are then described. The last section concludes the chapter presenting the research development plan.

3.1 Core Idea and Motivation

What affects operators' actions when observing remote events on a visual display, is their visual comprehension of the observed scene. This impacts effectiveness of monitoring actions and decision-making. The role of VR visual experience is then of paramount importance and worth being studied and assessed.

When studying VR visual experience, a key role is played by the system's provided sense of immersion. This profoundly affects user' sensation, as it makes users feel present in the observed virtual world. The resulting sense of telepresence greatly influences user's behaviour, and performance. What we wish to understand and assess can then be called VR immersive visual experience.

Understanding the VR immersive visual experience calls first of all for experiencing it. This means that a VR system needs to be tried out and one of more applications need to be considered.

- *VR System*. As introduced in previous chapter, several VR systems have been proposed in the last decades, featuring different display technologies. There is today a display system that has greatly evolved and has become the prevalent VR display. This is the Head Mounted Display (HMD) system.
- *Applications*. VR technologies can be adopted on a wide range of application, which has made the use of VR nearly ubiquitous. The resulting experience may nonetheless be very different based on the specific applications we are considering. This in turn means that the target application and consequently the type actions it calls for play a great role and may lead to different user experience.

The above two elements clearly indicate we need to identify VR systems and applications we want this project to target and use in our experiments. The chosen VR System is the HMD. As for applications, we consider them based on their practical relevance, and most importantly we look at actions a chosen applications call for. Of particular interest are actions of observation, monitoring, and positioning of observed objects. They are typical actions in many applications related to teleoperation, including: tele-exploration, tele-monitoring and tele-intervention. In particular, the following type of actions are considered.

- *Overview, monitoring and exploration.* The focus is on system output and visualisation, towards a more natural and intuitive visual comprehension of events and activities. Overview represents gaining general situational awareness of the observed objects, whereas monitoring represents tracking the objects' pose (position and orientation). Exploration includes the ability of moving freely around within the observed environment, while looking at all presented objects. Overview, monitoring, and exploration contribute to decision making.
- *User's Interaction.* The focus is on a more effective visual response when the user interacts with the current scenes, to achieve more informed insights and further speed up decision-making. Interaction considers user's movement as well as input and output devices.

The above listed type of actions represent activities where immersive display systems can bring significant advantages compared to traditional displays. E.g. observing through an HMD an object flying around us can make a user understanding the object position and direction with a much higher accuracy than observing the same flying object through a flat monitor display.

The three elements an HMD display possess, which we identify as key factors are: 360-degree field of regard (defined as how much of the user's view that is taken up with a computer displayed view [22]), S3D viewing, and isolation from the user surrounding environment. Concerning user-interaction with the observed world, a key factor is head-rotation. This allows users to observe a discover the surround environment by rotating their head, i.e., the same way we naturally see and discover the world around us, in our everyday life experience.

The next sections introduce and motivate each of the three proposed development phases. They are:

- *Phase 1: The Immersion Advantage;*
- *Phase 2: The Photograph Advantage;*
- *Phase 3: The True-Dimension Advantage*.

3.2 Phase 1: The Immersion Advantage

The considerations made in the previous section delimited our first phase of study, development and experimentation, which aimed at understanding and assessing the advantage of using an immersive display such as the HMD, in applications of command and control and pilot training.

In this type of applications, the use of immersive technologies can lead to greater comprehension and accuracy of observation, making the decision process faster and more effective. However, we can ascertain that the spread of immersive technologies has shown slow take-up, with VR adoption limited to few applications. This appears due to the often-necessary change of operating paradigm and some scepticism towards the "VR advantage". There is, therefore, a need to assess the benefits a VR system can bring to decision-making within commercial applications and the contribution immersion brings to user performance.

This will help system designers understand when immersive technologies can be proposed to replace or complement standard display systems. A desktop monitor may, for example, be more suitable for visual overviews, whereas a VR headset can serve best on focused observations and actions.

The research phase one proposes an evaluation of the advantage of using a VR headset compared to a standard desktop monitor. Three user studies are conducted in this phase.

The first user-study considers actions of identification, discovery, positioning and cognitive learning (search and find, routine movements and teleoperation) within pilot-training operations. It focuses on assessing the role of HMD's immersive and 3D viewing, looking at the role of HMD's immersive elements such as: isolation and 360-views.

The second user-study considers actions of monitoring and overviewing when observing through an immersive display. It focuses on assessing user-scene interaction with two different user controllers.

Concerning environment representation, the decision has been for the phase one to adopt what is currently the mainstream solution for VR visual representation. I.e., to use of computer graphics for the rendering of the visual scenes. Such choice is relevant to understand current VR visual implementations and to assess the advantages and limitations the use of VR brings on mainstream visual representations. It has also been considered a good base for the project second phase where we aim at understanding and assessing photo-based visual representation of VR environments.

The research phase aims at building new knowledge by evaluating findings and identifying advantages in complex scenarios. The research relies on analysing both objective measurements (time and accuracy), and subjective ratings related to human factors (sense of presence and depth perception).

3.3 Phase 2: The Photograph Advantage

For the second phase of research, the focus was on studying, developing and experimenting with photo-based visual representation of VR environments. This is an innovative way of developing VR environments representation, which provides great advantages in terms of visual realism and sense presence.

Photo-based VR is a solution of easy implementation, which represents a very convenient way to visualize real remote location and interact with them. It is of great interest because of the development and spread of 360 cameras featuring low cost and high-resolution. Photo-based omnidirectional imaging has become directly exploitable for VR, with their combination proven suitable for: remote visits and realistic scene reconstruction, operator's training and control panels, surveillance and e-tourism.

Photo-based VR is very interesting for research because beside the advantages (low-cost and fast-capture) it has also limitations in terms of vantage points and scene navigation. Furthermore, there is however a limited amount of scientific work assessing VR experience and user's performance in photo-based environment representations.

Concerning environment representation, we still rely on the use of computer graphics for the rendering of the visual scenes. However, the graphical environment is simple, because it relies on a model made by either a sphere or cube, where photograph is mapped onto. It then becomes very relevant the way the photo-mapping is implemented, and there are today new methods being developed, while user's interaction with the scene is managed by the graphic system, which can also manage some (additional) lighting effect.

This type of representation gives great relevance to the captured photograph (video can also be used), therefore scene illumination, image deformation and illustrated depth characteristics all play an important role. Visual realism is one of the strongholds the photo-based visualization brings to VR visual experience, which therefore becomes a major element we want to assess in this study.

Unfortunately, photo-based VR increases the complexity of image formation and display in terms of system parameters. A camera system, possibly on stereo-viewing setup, now needs to be considered in addition to the display system. Image processing can also play a role, so as the graphical visual rendering. This results in a delimitation of the investigated aspects, while focus remains on the impact of imaging.

The research phase two proposes assessing the effect of photographic realism in VR when observing real places through a VR headset, for two different pixel-densities of the screen display, two different types of location environment and two different levels of sense of place. This results in two user studies (user studies 3 and 4). The comparison relies on the observation of static three-dimensional and omnidirectional photorealistic views of environments.

The goal of this phase is to gain an insight about how photographic texture can affect perceived realness, sense of presence and provoked emotions, as well as perception of image-lighting. The two user studies are conducted based on subjective rating given by users to a number of display and human factors.

3.4 Phase 3: The True-Dimension Advantage

Photo-based VR representations observed through a VR headset paves the way for user to perceive the actual space dimension. This is referred as true-dimensional visualization.

A specific aspect the project studies and evaluates, with the aim of understanding its mechanics, while also developing a novel solution, is related to depth perception in photo-based VR representations.

There are a wide number of applications in VR, both relying on photo-based rendering and traditional graphical rendering, which would benefit for realistic portrayal of distances.

This is of particular interest when using HMDs, because the observed environments can be explored the way we naturally explore the world, i.e., through head rotation.

This in turns means that there is no longer a need to use unrealistic wide FOVs to observe the represented environment. This, which was often a need, e.g. when observing a wide room on desktop display, and led to depth mis-judgment and erroneous distance estimation, is no longer required when observing through an HMDs. Users wearing an HMD can turn their head around to observe and discover the environment around them, which is also the natural way humans observe objects and places in their everyday life experience.

The user’s observed FOV can then be made similar to that human naturally experience when looking at the real world with necked eyes. This affects realistic size observations.

The research phase three proposes assessing depth perception and distance estimation when observing real places through a VR headset, for different environment types. The proposed comparison relies on the observation of static three-dimensional and omnidirectional photorealistic views of environments. The purpose is to gain an insight about how photographic texture in VR scenes observed through an HMD, can affect the perceived space dimension.

The proposed user study (user study 5) assesses depth perception through subjectively ratings depth impression and contribution and then measures errors in distance estimation too.

A number of focused assessments are then devised and conducted to analyse the role of hardware and software FOVs and to develop a new solution that correctly re-adjust software FOV during navigation (Concentric Sphere method).

Finally, a new user study is proposed to verify HMD typical underestimation of perceived distances in VR [23], and to devise a solution that reduces or counter-balances the underestimation of distances.

3.5 Research Development Plan

The described research objectives were proposed to be achieved through three main phases where the advantages of VR immersion, photo-based VR and true-dimensional visualization were studied and assessed. Those phases also included software developments.

The overall research development included three main stages, which were roughly associated to each year of study. They were:

- *Stage 1: Learn, Experience and Develop;*
- *Stage 2: Enhance Visual Realism and True-Dimensional Visualization;*
- *Stage 3: Assess VR Enhanced Telepresence and Visual Comprehension.*

Stage 1: Learn, Experience and Develop

The first stage was dedicated to learning about the subject and application areas, and to experience the relevant software and hardware. This stage would then represent a first necessary step to investigate the advantage brought to VR viewing by greater visual realism and more correct perception of observed space dimension. This stage therefore focused on the following objectives:

- Comprehend virtual reality technologies and the main associated characteristics and parameters, including system parameters and human factors.
- Familiarize and experiment with latest VR technologies.
- Comprehend main elements characterizing VR visual realism, photo-based VR and visual comprehension.
- Analyze the role of camera and display parameters and their relation to depth perception, presence and comfort.
- Comprehend and synthesize appropriate graphical tools allowing for designing VR visual interfaces and graphics. Develop software to experiment new solutions.

Associated to the above objectives there were a number tasks, which were grouped in the three implementation steps described in the Figure 18.

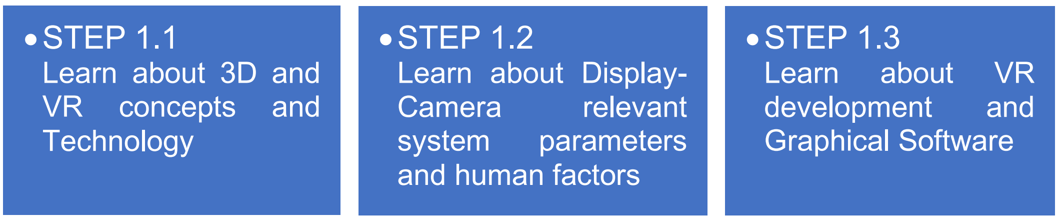


Figure 18 – Steps Stage 1: Learn, Experience and Develop.

In this type of research in the engineering field is deemed relevant to start by exploring virtual reality technologies, their advantages and limitations, potential applications, and gain deeper knowledge on the parameters governing the relevant hardware systems while also start reviewing the related literature. The studies also planned to include learning how to develop a VR project, the different associated libraries, tools, and gaming engines. The focus was expected to be on learning the Unity3D software and how it works and develops on Head Mounted Display systems such as Oculus and HTC Vive. The programming activity also were to include system interaction through their controllers and the streaming of different cameras in VR.

Stage 2: Enhance Visual Realism and True-Dimensional Visualization

The second stage was dedicated to study an enhancement of visual realism by experiencing and assessing the involved parameters and therefore learning from the experiments' outcome. It focused on the following objectives.

- Advanced synthesis of appropriate graphical tools allowing for designing VR visual interfaces and graphics. Develop software to assess new solutions.

- Design and experiment new solutions to enhance visual realism based on three and true dimensional visualization of remote images.
- Synthesize obtained outcome within scientific publications.

The design of new solutions was expected to be based on small steps of innovation. One often needs to bring improvements into small visualization steps because they may all contribute to a more correct dimensional perception. To achieve the above, one needs to assess the current status of a solution and then improve it by bringing some innovation for a more effective outcome.

Associated to the above objectives there were a number tasks, which were grouped in the two implementation steps described in the Figure 19.

Figure 19 - Steps Stage 2: Enhance Visual Realism and True-Dimensional Visualization.

The proposed type of research was expected to naturally lead to produce a number of scientific publications based on the performed experiments. The publications were typical to target international conferences focusing on specific assessment.

Stage 3: Assess VR Enhanced Telepresence and Visual Comprehension

The third stage included the assessment of both VR technologies and the enhanced telepresence and visual comprehension. It focused on the following objectives B1, B2, B3, C1, C2 and C3.

- Comprehend assessment of immersive and non-immersive technology solutions, including system and usability evaluations.
- Comprehend experiment design, conduct pilot and formal evaluations assessing advantages and limitations of VR visual interfaces. Evaluate results obtained from experimentation.
- Synthesize obtained outcome within scientific publications.
- Synthesize assessment of VR Technology and True-Dimensional Visual Realism
- Evaluate obtained results.
- Synthesize thesis manuscript and scientific publications.

Associated to the above objectives there were a number tasks, which were grouped in the two implementation steps described in the Figure 20.

- **STEP 3.1**
  VR Technology Assessment: Immersion, Interaction, Action Training, Navigation
- **STEP 3.2**
  Assess VR Technology and Realistic Viewing

Figure 20 - Steps Stage 3: Assess VR Enhanced Telepresence and Visual Comprehension.

The proposed type of research was expected to lead to a more comprehensive study overview in the field while proposing innovative solution. It was then expected to produce more scientific publications out of the performed studies, hopefully targeting good ranked international journals.

Yearly Progress Visual Description

The accomplished tasks are visually described within the diagram in the Figure 21, where progress can be seen also in relation to the three years of study. The coordinated progress of the specific topics has been a key element of the PhD project development.

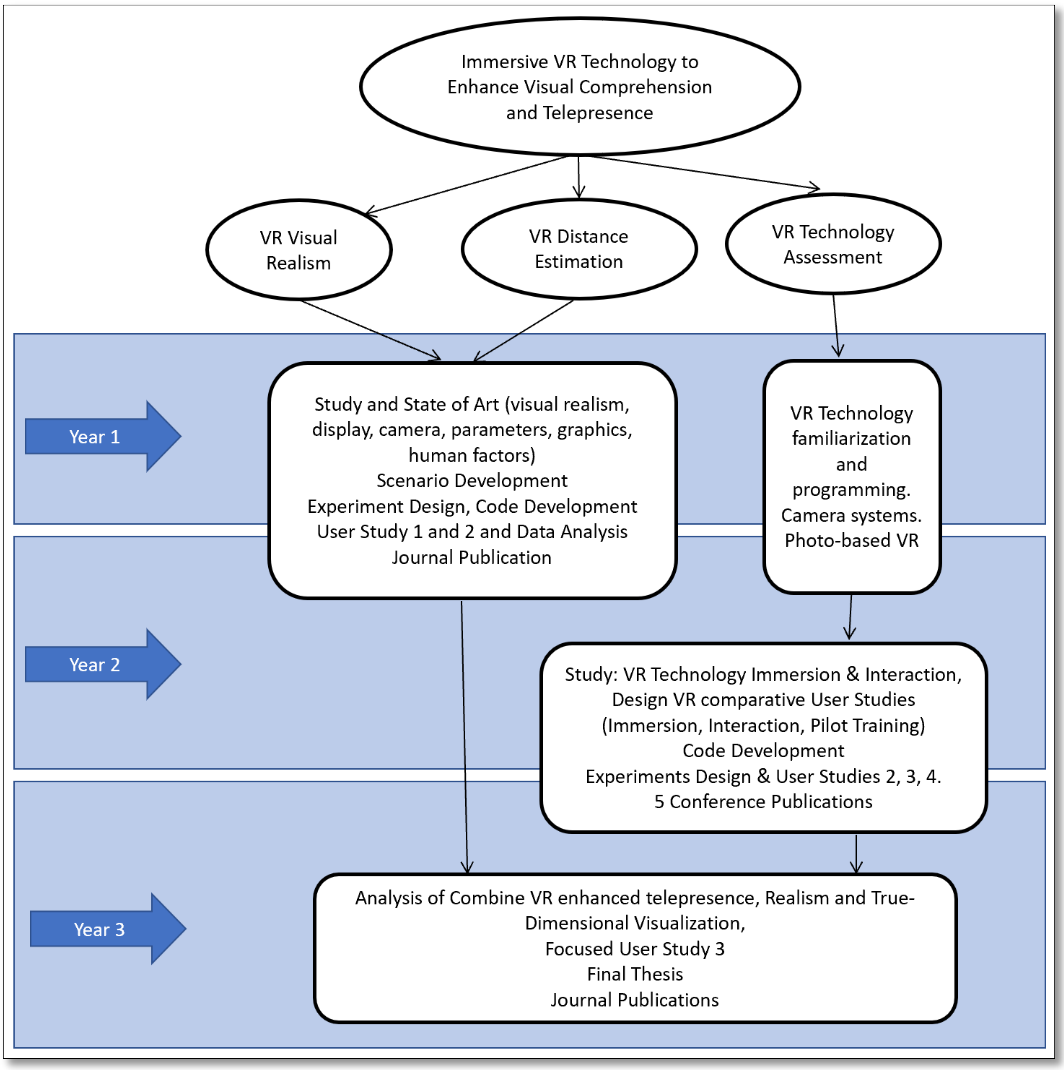


Figure 21 - Diagram of the accomplished tasks in relation to the three years of study.

4. The Immersion Advantage: Display and Interaction

This chapter presents two user studies that were performed to gain insight into the role of immersion in VR visual experience. We have learnt from the literature that the performance of a VR systems may largely depend on the applications. We have therefore considered a representative training application, which addresses different key tasks (search and find, procedure learning, and teleoperation) and a key user interaction aspect. The following sections present the conducted studies including the related state of the art, research questions, experiment design and execution, results analysis and conclusions.

4.1 User Study 1: Pilot-Training

The growing demand for aircraft pilots and spread of drones has resulted in a greater need for pilot training and cost reduction. Flight simulators have therefore become increasing popular and regularly used for training new pilots and maintaining knowledge of experienced aviators. New VR technologies have been proposed with the promise of speeding up pilot training. Immersion is a key aspect, well supported by today's systems featuring high-quality displays, omnidirectional viewing and multi-point audio. Haptic feedback is also being introduced by researchers but its application is limited to vibrations.

The high-level of immersion VR systems brings higher realism to training, narrowing the gap between simulation and face-to-face training. However, most students and trainers feeling comfortable just using traditional simulators and do not necessarily want to change the way they learn and teach. There is nonetheless new interest towards VR solutions. Different air forces, e.g. those in the USA, have introduced the use of VR technology in their education programs [24], while literature works have shown numerous experiments proving VR technology has the potential to produce pilots faster than other methods and cannot be avoided [25]. There is therefore potential for VR simulators for adoption, replacement or in combination with traditional simulators.

In case of vehicle teleoperation, VR technology brings drivers close to being in the actual driving seat, therefore even beyond what is actually feasible when remotely operating vehicles [26]. Nowadays we deem both VR systems and 2D-screen monitors to be considered for training, as both have advantages. E.g. students can start with the traditional method of sitting in a classroom and look at large displays to then follow-on to a focused VR headset based exercise. We need therefore to understand when the immersive experience provides additional value to training and when instead immersive experience is not needed or even undermining learning effectiveness. It is crucial to assess the usability of VR solutions to better understand their effectiveness.

In this section, we assess performance of aviation training tasks on both a 2D-screen monitors and a VR headset. The assessment includes two different types of training. One concerns flight deck training performed by a user while sitting in front of a cockpit. This has the aim of practicing with instrumentation and basic take-off preparation actions for take-off.

The other training concerns skilful tele-operation practice and consists of remotely flying a drone. This has the aim of learning vehicle tele-guiding skills in presence of time measurement and dynamic scenes. We gather several objective and subjective data to provide us insight on advantages and drawbacks of using VR headsets, which represents a base for future system designers.

4.1.1 State of the art

Works in the scientific literature addressing the "immersive advantage" include immersion concepts [27], telepresence [28], using comparing traditional computer desktop monitors to HMD to study brain activity and sickness [29] [30], or comparing to virtual displays and virtual interactions [31] [32]. The term immersion is also not used consistently, and terms such as presence and engagement are mentioned interchangeably. Immersion is typically described as a technology [27] that provides the capabilities of "shut out physical reality" by replacing the human experience of seeing, hearing and moving with an artificial experience. Egocentric perspective is often considered a critical aspect. On the other hand, presence mainly refers to the sensation experienced when virtual reality feels real.

There are different ways of measuring presence varying from questionnaires, subconscious body reactions, intentional physical movements to interviews [12]. Typically, presence is studied with other aspects to provide information to future researchers to influence how applications are designed [33].

A few aspects that are typically paired together include accuracy, time and satisfaction. Bueckle et al. [31] provided a way of measuring human body organ size, position and orientation using VR compared to Desktop at different difficulty levels. VR users were three times as fast in rotation and a third more accurate than desktop. However, positioning across VR and Desktop conditions had no significant differences.

This positioning and other skills rely on S3D performance that relies on a range of visual functions from contrast sensitivity, visual crowding and visual attention that can be trained [34] that improves stereo acuity that contributes to a better depth discrimination. However, researchers only measure discrimination on depth discrimination distances of less than 10 meters when viewing S3D environments. Naceri et al. [35] compared a real to a virtual environment at distances varying between 1.4 - 2.4 meters and found VR distance estimations less precise, but still significant.

If S3D is reviewed more generally, there are several types of tasks researchers attempt to measure include distance, finding/identifying/classifying, manipulation of real/virtual objects, navigation, spatial understanding/memory/recall, learning/training/planning [36].

The concept of immersion or presence has been around for many years. However, there is some confusion due to varying definitions from a) hardware fidelity, b) psychological aspects that could consist of what the person sees, or, what the person experiences from the virtual environment [37]. Furthermore, this immersion can help or hinder performance [38], [39].

Accurate and time-critical decision-making is a topic of paramount importance when piloting an aircraft. Coupled with this is finding the right instrument on a control panel and being able to operate it promptly. Visual search is a fundamental aspect in which efficiency depends on the amount of information presented to a user. What may hinder visual search efficiency is the presence of “distractors”. They represent objects and features present in the user’s visual field that is not of interest to the user. In other words, they represent obstacles in finding the target object.

Emani et al. [40] discussed the effect of “visual distractors” and how these affect cognitive load. They conclude that any visual stimuli that were not relevant to the task significantly increase the objective measure of cognitive load on the parietal channels (alpha and theta brain waves) using EEG-BCI technology. Furthermore, the linear relationship was strongest for the lowest performing group where subjective cognitive load was greatest then there is a decrease in BCI performance.

Olk et al. [41] investigates measuring visual search in immersive VR and on a 2D monitor. In this work, users searched for an item while presenting several familiar distractors. At times objects are placed in the scene that are similar or dissimilar, to understand how it effects user performance. Reaction time is slower when placing an object that is dissimilar. This slower reaction time could be because the objects are familiar to the user, despite the objects being a distraction. The same result exists both in the VR and 2D monitor conditions.

The effect of using different display settings in an AR system is investigated by Marquard et al. [25]. This paper compares limited Field of View to audio-tactile approaches on situational awareness. It discovered that the visual approach is the fastest, but the audio approach provided the best improvement on situational awareness due to the focus on relying more on memory.

Xu et al. [42] studied reaction time when examining similar and different symbols to understand the cognitive processes involved to aim to explain user performance. Researchers found that symbols near each other interact in some way affects performance. Secondly the processes that affect homogeneous search also affect heterogeneous searches. It is believed that search interactions occur when similar objects are placed next to each other distracts the user and slows down their performance, this is supported by previous research [43].

Walters and Walton [44] focused on pilot training. It underlines the lack of literature works addressing links between search task performance, collaboration and sense of presence when using VR systems. Makranski and Peterson [45] addressed learning in an immersive environment. It presented a conceptual model by synthesizing existing literature and demonstrating the model factors that lead to learning in an immersive environment. There is a link between learning in a traditional manner and an immersive environment. The taxonomy suggests that the technological factors match affordances, cognitive factors and learning outcomes.

Connected to learning are memorization approaches. Memory palaces are among the most popular types of traditional memorization that also applies to VR. Krokos et al. [46]

uses virtual reality to look around a spatial environment to test recall compared to a desktop environment. It discovered that immersion aids recall within a specific layout within a familiar indoor environment. Staton et al. [47] argue people use the layout as a means of remembering.

Memorization effectiveness depends on the technique and the approach. Changing one aspect of the experience can affect performance that relies on memory and learning. Johnson-Glenberg et al. [48] compared 2D desktop and 3D virtual reality STEM environment with different levels of embodiment and test time. Embodiment is defined as watching or interacting. There is significant proof for a high-level embodiment affects learning, with the VR group performing the best.

4.1.2 Proposed Investigation and Experiment Design

The study aims to evaluate the advantages of using immersive and non-immersive display technologies in tasks of identification, discovery, positioning and cognitive learning (search and find, routine movements and teleoperation) within pilot-training operations.

The assessment focuses on assessing the role of HMD's immersive and 3D viewing by comparing performance between what is currently the most used display system, i.e., the desktop monitor (DM), and a recent HMD.

4.1.2.1 Research Questions

The research question involves the following aspect.

- *DM vs HMD.* Does operating through an HMD improve, identification, monitoring, positioning and cognitive learning performance?

The study focuses on assessing the performance advantage of using a VR headset for flight deck training (FDT). Test-users are asked to execute two typical training procedures while their performance is recorded. Skilful tele-operation practice (STP) is also assessed for constrained flying. User performance is evaluated for time-completion and precision of operation. Human factors are also taken into consideration. In particular, we look at presence, comfort and ease of use.

4.1.2.2 Evaluation Measurements

Literature works have shown immersive systems can lead to higher spatial comprehension [31]. This is an aspect pilots heavily rely on both during training at cockpit and aerial vehicle teleoperation. We aim therefore at replicating real training actions that rely on spatial comprehension, on both desktop monitor (DM) and VR headset (also known as head mounted display - HMD). The data gathered during experimentations include:

Objective measurements:

- *Time-Completion*. Time elapsed to complete a required training action. This typically include object identification and response time (FDT) and navigation time (STP).

- *Precision*. Correctness of the applied procedure (FDT) and vehicle teleoperation (STP).

Subjective ratings:

- *Presence*. The perceived sense of being there (FDT - sitting at vehicle's cockpit).
- *Comfort*. General body and eye comfort during simulation action.
- *Ease of Use*. System usability and ease of operation.

4.1.2.3 Tasks and Procedure

The FDT experiment includes the following tasks:

1. *Search & Find (S&F)*. Users are asked to search for (three) specific instruments at cockpit's dashboard and indicate their positions with pointing with their finger. This works also when wearing the HMD as users see a virtual replica of their hands and fingers, allowing them to naturally point to instruments. Time-completion and precision of operation are recorded. Figure 22 shows the cockpit dashboard (based on Cessna 152).

Figure 22 - The cockpit dasboard (Cessna 152).

2. *Check List*. Users are asked to perform a specific check list procedure at cockpit. It consists of: (1) switch activation; (2) yoke push & pull; (3) yoke turn-left & turn-right; (4) throttle push-pull; (5) switch deactivation. Users are told about the procedure initially. They then need to remember and replicate the actions. Time-completion is recorded while precision of operation is marked by the test monitor on a 3-point scale after each operation. A score of 1 means poor performance and a score of 3 means best performance. Figure 23 shows a test-user during S&F on DM.

Figure 23 – Left: test users during S&F on DM. Right: our HMD system.

The STP experiment includes the following tasks:

1. *Tele-Guide*. Users are asked to tele-operate the drone inside a narrow tunnel avoiding collisions with its walls. Elapsed time is measured when the experience starts and when the drone exits the tunnel. Precision of operation is based on time employed to disengage after a collision. Figure 24 shows a DM and VR tunnel view and a test-user during teleguide.

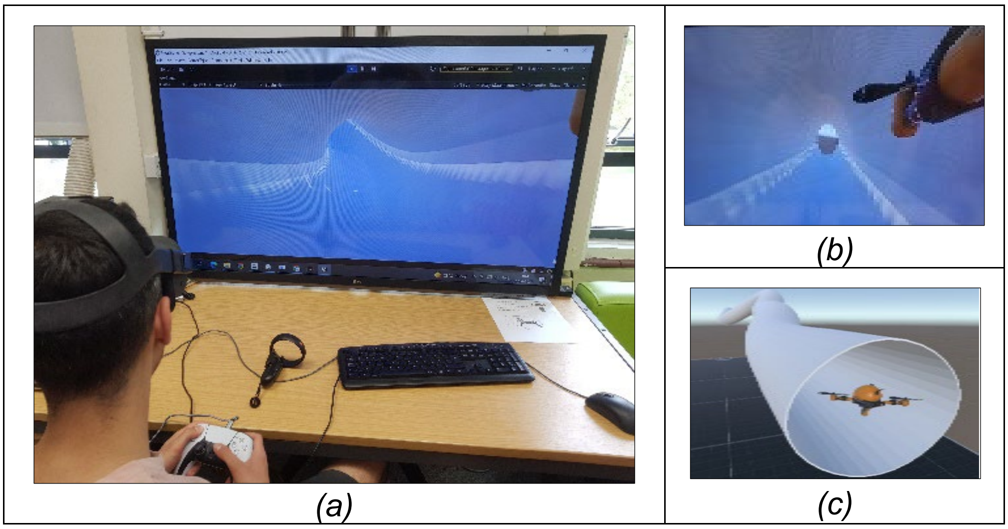


Figure 24 – (a) Test users during Tele-Guide; (b) An example tunnel view; (c) Our tunnel where the drone runs.

4.1.2.4 System Apparatus

The hardware used in our experiments includes high-specs PC with *nVidia* GPU GeForce RTX 3080Ti graphic card. The software used is *Unity*. The VR headset is an *Oculus Quest 2*. Driving and switch activation/deactivation is achieved through a console-game controller (*Sony PS5*).

4.1.2.5 Experiment System and Design

Twelve participants conducted a within-subjects' evaluation over the two systems. Participants had none to good experience with VR systems and computer games (uniformly distributed). Their age ranged between 20 and 55 (average of 27.4 years old), and were 53.3% females. The testing scenario and system sequence were assigned according to a predetermined counterbalanced schedule to avoid fatigue and learning effects. Our test was designed following literature recommendations and conformed to traditional approaches in terms of forms and questionnaires [49]. Initially, we provided participants with information, consent form and pre-test screening. A system practice session was then administrated followed by two repetitions of the operation session. Questionnaires were then administrated. The STP procedure has introduction, practice, flight task, then questionnaires. The flight task and questionnaires are done once for Desktop and once for VR. Figure 25 summarizes the procedures for FDT.

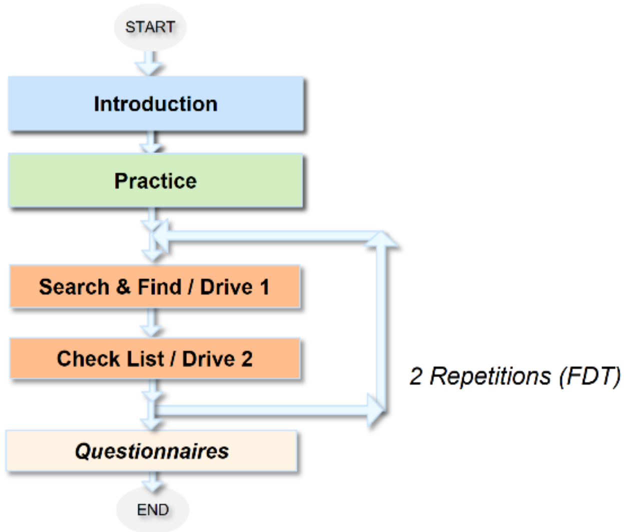


Figure 25 - Testing procedure.

We chose to rely on standard questionnaires. They were: *Presence* (Igroup Presence Questionnaire - IPQ) [50] assessing: general presence (Q1), spatial presence (Q2-Q6), involvement (Q7-Q10) and realism (Q11-Q14); *Comfort* (Simulator Sickness Questionnaire - SSQ) [51] [52]; *Ease of Use* (System Usability Questionnaire - SUS) [53]. Answers were provided on a 7-point Likert scale, except for the SUS (5-point) and SSQ (7-point). We computed mean values of acquired data and measured statistical significance of results by estimating the Student's t-distribution for paired comparison with repeated measures. We set p-value= 0.005 as threshold for significance. We also estimated the standard error of the mean (SE) for each comparison. The results are presented and analysed in the next two sections according to the research questions.

4.1.3 Results and Analysis

4.1.3.1 Time and Precision

Table 1 shows results. Users were considerably faster in searching and finding instruments in the Cessna dashboard when observing through the HMD compared to DM (64% less time in average). The HMD performed significantly higher than DM (p=0.047). Precision of operation was also significantly higher (p=0.046). Users employed nearly the same time to perform the check list, with a significant higher precision on the DM (*p*=0.022). The outcome of the assessment is very clear. The HMD shows great effectiveness in more challenging tasks such as S&F, whereas tasks easier to perform as the Check-List, which mainly relied on action memorization, are done better on DM.

Drone navigation did not show significant time difference between DM and HMD, which is opposite to what we expected based on the VR teleoperation literature. Typically, wide-angle immersive views call users to spend more time to carefully observing the surrounding environment. It was the simplicity of our environment and tunnel texture that made it less attractive for our users to look around.

Drone teleoperation through the HMD scored nonetheless 55% more precise, which is a clear indication of HMD potential.

	Search & Find		***Check List***		***Drone Navigation***	
	Time	**Precision**	**Time**	**Precision**	**Time**	**Precision**
DM	29.069	2.083	12.864	2.667	27.578	4.919
HMD	19.322	2.583	13.080	2.083	28.467	2.265
DM-HMD	**0.047**	**0.046**	0.447	**0.022**	0.383	0.282

Table 1 – Shows the elapse time and precision (how correct a user is at finding an object, getting the checklist correct, or avoiding collisions)

4.1.3.2 Presence

Figure 26 shows results. The HMD scored higher in most of presence indicators performing significantly better on spatial presence (Q5 *p*=0.006) and involvement (Q7 *p*=0.066), (Q9 *p*=0.060). The HMD characteristics in terms of isolation, field of regard (FOR) and stereoscopic 3D visualization (S3D) are definitely helpful to comprehend cockpit dashboard instruments and the tunnel area. Presence definitely plays a role on S&F tasks precision and timing. Presence questions are shown in a Table 2.

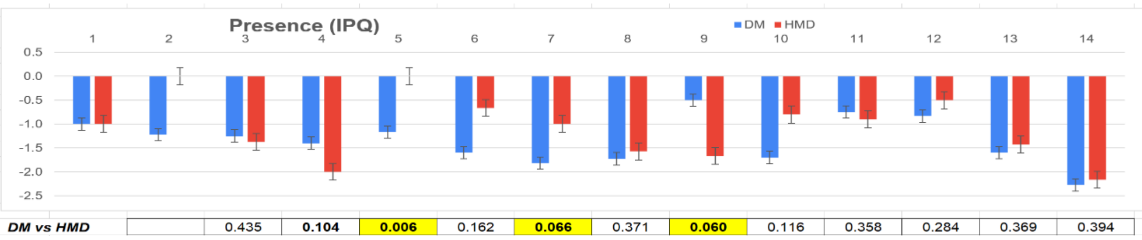


Figure 26 - Mean, SE and Student’s p-values for Presence.

Presence	From (-3) to (+3)
1) In the computer generated world I had a sense of "being there".	Not at all -- Very much
2) Somehow I felt that the virtual world surrounded me.	Fully disagree -- Fully agree
3) I felt like I was just perceiving pictures.	Fully disagree -- Fully agree
4) I did not feel present in the virtual space.	Did not feel -- Felt present
5) I had a sense of acting in the virtual space, rather than operating something from outside.	Fully disagree -- Fully agree
6) I felt present in the virtual space.	Fully disagree -- Fully agree
7) How aware were you of the real world surrounding while navigating in the virtual world? (i.e. sounds, room temperature, other people, etc.)?	Extremely aware -- Moderately aware -- Not aware at all
8) I was not aware of my real environment.	Fully disagree -- Fully agree
9) I still paid attention to the real environment.	Fully disagree -- Fully agree
10) I was completely captivated by the virtual world.	Fully disagree -- Fully agree
11) How real did the virtual world seem to you?	Completely real -- Not real at all
12) How much did your experience in the virtual environment seem consistent with your real world experience ?	Not consistent -- Moderately consistent -- Very consistent
13) How real did the virtual world seem to you?	About as real as an imagined world -- Indistinguishable from the real world
14) The virtual world seemed more realistic than the real world.	Fully disagree -- Fully agree

Table 2 – Presence Questionnaire

4.1.3.3 Comfort

Figure 27 shows results. Comfort is a relevant indicator in VR-based pilot training because of the extended sessions. The scores for comfort were very low on both systems which is very reassuring. The HMD typically bringing more discomfort, which is in line with the literature. Comfort questions are shown in a Table 3.

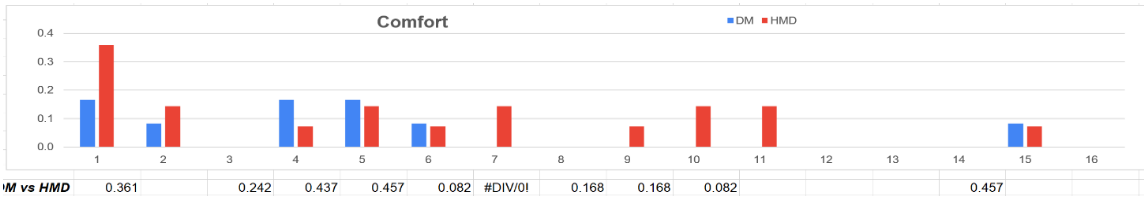


Figure 27 – Mean, SE and Student's p-values for Comfort.

Comfort	4 choices			
1) General discomfort	None	Slight	Moderate	Severe
2) Fatigue	None	Slight	Moderate	Severe
3) Headache	None	Slight	Moderate	Severe
4) Eye strain	None	Slight	Moderate	Severe
5) Difficulty focusing	None	Slight	Moderate	Severe
6) Salivation increasing	None	Slight	Moderate	Severe
7) Sweating	None	Slight	Moderate	Severe
8) Nausea	None	Slight	Moderate	Severe
9) Difficulty concentrating	None	Slight	Moderate	Severe
10) «Fullness of the Head»	None	Slight	Moderate	Severe
11) Blurred vision	None	Slight	Moderate	Severe
12) Dizziness with eyes open	None	Slight	Moderate	Severe
13) Dizziness with eyes closed	None	Slight	Moderate	Severe
14) Vertigo (It is experienced as loss of orientation with respect to vertical upright)	None	Slight	Moderate	Severe
15) Stomach awareness (It is usually used to indicate a feeling of discomfort which is just short of nausea)	None	Slight	Moderate	Severe
16) Burping	None	Slight	Moderate	Severe

Table 3 - Comfort Questionnaire.

4.1.3.4 Ease of Use

Figure 28 shows both systems have a high mean that represents the system is appreciated and user friendly. They also score high as well-developed product and quick to learn. The HMD scored generally higher than DM as mean value (except for Q8), and VR scored significantly easier to use than DM (Q3 p=0.027). HMD also scored significantly higher regarding ease of learn than DM (Q7 p=0.041). Comfort questions are shown in a Table 4.

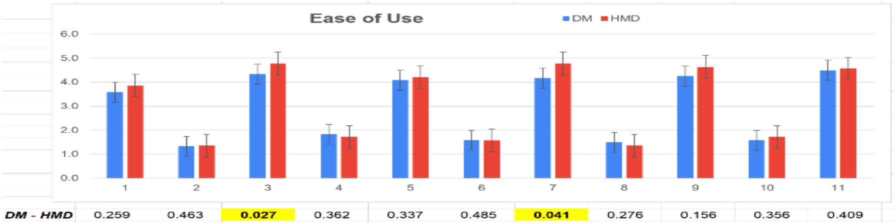


Figure 28 – Mean and Student's p-values for Ease of Use.

Ease of Use	From (1) to (5)
1) I think that I would like to use this product frequently.	Strongly Disagree -- Strongly Agree
2) I found the product unnecessarily complex.	Strongly Disagree -- Strongly Agree
3) I thought the product was easy to use.	Strongly Disagree -- Strongly Agree
4) I think that I would need the support of a technical person to be able to use this product.	Strongly Disagree -- Strongly Agree
5) I found the various functions in the product were well integrated.	Strongly Disagree -- Strongly Agree
6) I thought there was too much inconsistency in this product.	Strongly Disagree -- Strongly Agree
7) I imagine that most people would learn to use this product very quickly.	Strongly Disagree -- Strongly Agree
8) I found the product very awkward to use.	Strongly Disagree -- Strongly Agree
9) I felt very confident using the product.	Strongly Disagree -- Strongly Agree
10) I needed to learn a lot of things before I could get going with this product.	Strongly Disagree -- Strongly Agree
11) Overall, I would rate the user-friendliness of this product as:	Worst Imaginable - Awful - Poor - Ok Good - Excellent -Best Imaginable

Table 4 - Ease of use questionnaire.

4.2 User Study 2: User-Scene Interaction

Interaction is a crucial aspect when operating in VR environments. It affects comprehension of the observed scene and its evolution, often referred to as situation awareness, with consequences regarding the user's reaction to observed events and decision-making.

When operating through VR technologies, the interaction involves both sensing user's action and providing commands. The technology for sensing users' actions typically includes tracking movements, e.g. head, hand and eye positions, and, more recently, voice commands. The technology for providing commands typically includes pushing buttons on keyboards or controllers and moving joysticks, while hand gestures have become more popular, so is to read eye movements.

Situation awareness has been investigated in various studies addressing interaction [54]. Prior work in interaction attempts to understand different aspects. These include understanding controllers that are either different or have the same changing visual elements [55], [56], [57]; analyzing movements such as head, body or eye movement [58] [59] [60]; comprehending ergonomics and VR sickness [30]; understanding interaction role by task performance [61], [27]; and understanding connection between performance or interaction [62].

When assessing interaction technology with VR, one needs to consider the specific application as this affects the required level for users' performance. For example, in command and control operations there is a need to overview scenes and events, perform focused observations, and act and provide commands. Very few works in the literature focus on interaction related to defense analysis, and nearly no literature to our knowledge in large-scale remote battlefields and VR headsets, besides Zocco et al. [63] Augmented reality work.

The use of HMDs, which has grown in the last years to become the dominant VR technology [64], can undoubtedly help achieve higher user presence and involvement because of the more heightened sense of isolation and immersion provided. Furthermore, it allows users to discover and explore the surrounding environment through the simple and natural gesture of head rotation.

The action of head rotation when observing through a VR headset is a very natural action used to discover and explore the surrounding environment. However, this action is not necessarily the most suitable for all types of observations in VR. We know, for example, that overviewing is often better dealt with when limiting body movements, as this helps reduce cognitive load and cyber-sickness. However, it is unclear how to best use interaction tools for user-scene observation, which often leads system designers to opt for traditional desktop-monitor-based solutions. Nonetheless, the advantage of immersive visualization and natural observation actions may be relevant for exploitation.

In this study we compare the use of HMD with observation viewpoint changed through controller's joystick (HMD-J), with the use of HMD with observation viewpoint changed through head rotation (HMD-H). Our application context is command and control operations within military defense, where users must gain high situational awareness in the observed scenarios, recognize threats accurately, and respond promptly.

4.2.1 State of the art

One of the affordances of virtual reality is movement can play a part in learning. The psychological aspects of learning and memorization are relevant for completing various tasks using movement within VR. For example, researchers are increasingly exploring sensorimotor experiences to understand their contribution to learning. Ratcliffe and Tokarchuk [65] reviewed 14 experimental sensorimotor studies. The review supplied positive learning outcomes using different approaches while discussing problems and the lack of methods for sensorimotor engagement within Virtual Reality.

Paired with the psychological aspects are the physical aspects of movement. Penumudi et al. [66] explored the effect of location on performance and physical load and discomfort, revealing that neck movement is more optimized for looking left and right and not up and down, as validated by subject discomfort measurement. In another study, Sargarnum et al. [59] investigated the concept of guided head rotation and amplified head rotation with VR viewing. The concept is to reduce physical strain by avoiding large head movements. The authors compared natural head rotation in VR to amplified head rotation with gamers and non-gamers. They discovered that gamers perform better for their preference and lower sickness when amplifying movement.

In another study by Christou et al. [60], gaze-directed and pointing motions are compared in the control of navigation tasks. In their evaluation, both head-rotation and eye-tracking are tested during the evaluation to assess the effectiveness of gazing alone while static or separating gazing from movement. Results showed control separated from gazing provided the least errors.

When the eyes move, they focus on a particular object using the accommodation and convergence process [67]. One use of this process is depth discrimination, which can take place with static binocular observations, but also with dynamic observation triggered by either head-rotation, observed-objects movements, or both. The concept of depth discrimination has been around for many years. However, much research is concerned with estimating distances [68] and less about using this depth discrimination or any underlying mechanics that go into decision making. Similarly, head rotation and locomotion affordance are studied to provide natural interaction with the virtual environment.

Depth discrimination is attained via human-vision binocular and monocular cues. These cues can be recreated in graphical rendering through stereoscopic 3D visualization and can be adjusted to provide different levels of depth impression. For example, Lamb et al. [69] used depth ranking, relative depth judgement and path tracing tasks in two stereoscopic rendering techniques (a) closest object, and (b) furthest object. The paper studied response time and error magnitude as objective measurements. As the static object focus is close, researchers discovered that performance response times are similar in complex path tracing tasks from one node to another, but the time taken is slower. Zocco et al. [70] presented a command-and-control study where a physical battlefield is represented as a virtual battlefield with symbols representing objects to identify risks in the virtual battlefield scenario. The author discovered that performance is faster using S3D viewing.

Focusing on a particular object could be more difficult in an immersive-based environment if the scene changes. Rahimi et al. [71] provided an experiment on automatic scene transitions and teleportation in Virtual Reality and explored the implications for spatial awareness and sickness. The study discussed switching between different types of automatic transitions from one viewport to another to understand its effectiveness while using an HMD. What is not known is the effectiveness of switching between different viewpoints in a manual effort using varying depth discrimination.

User movements and interaction also play a key role in involvement and, therefore, presence. Sharples et al. [30] explored two display conditions with active natural controls and two with passive controls, where the participant had no control of the virtual world. What is less understood is less natural active game controls, yet highly practiced depending on individual tendencies like gaming experience.

When operating through VR systems, head movements are closely related to comfort, which may concern visual fatigue, nausea, body tiredness and related psychophysics aspects. Therefore, a subjective rating is very common for assessing comfort, typically using weighted question sets of the Simulator Sickness Questionnaires (SSQ) [51]. In addition, some researchers advise discussing the factors in motion sickness and how to reduce or prevent it [72]. In contrast, other researchers seek to measure motion sickness predisposition through Motion Sickness Susceptibility Questionnaire (MSSQ) [73].

4.2.2 Proposed Investigation and Experiment Design

When observing through a VR headset, users typically have 360° views of the surrounding operational area, which needs to discovering through commands of interaction. Although viewpoint change by head rotation is intuitive and more natural than operating a joystick, it may affect comfort, e.g. in terms of cyber-sickness or neck pain, especially when such operation is continuous and rapid. Furthermore, monitoring actions include overviewing, which is normally considered more suitable for stable views. The above makes us wonder whether a joystick-based control would be more convenient for some applications, such as command and control, typical of military defense operations. This type of application calls for careful monitoring of actions and effective decision-making.

The study focuses on assessing the use of head-rotation compared to a joystick to trigger changes in observation viewpoint when wearing a HMD. The focus is on the effect on user performance (accuracy and time) and affected human factors (presence, comfort, ease of use and depth perception).

4.2.2.1 Evaluation Measurements

Literature works have shown immersive system can lead to higher accuracy in comprehension of remote events and tele-operation tasks [27]. However, some works have also demonstrated that immersive systems' wider views and 3D appearance can lead to more extended observations and action times [74]. Timely decision-making is a key aspect of C2 operations, we want to experiment C2 applications through the use of immersive systems and analyse their performance. Related human factors are also explored, in particular, sense of presence, depth impression, visual comfort and ease of use.

In summary, we gather data after the following six factors:

Objective measurements:

- *Accuracy*. Correct objects identification.
- *Time*. Elapsed time to complete monitoring and identification tasks.

Subjective ratings:

- *Presence*. The perceived sense of being there.

- *Comfort*. General body and eye discomfort.
- *Ease of Use*. System usability and ease of operation.
- *Depth Perception and Learning*. Delivered impression of 3D appearance, identification and memory recall.

4.2.2.2 Tasks and Procedure

Our experimental scenario includes two clusters of drones represented through graphical icons, which move towards the observer. Each cluster includes hostile and friendly drones (those hostiles are colored red). The users are asked to understand, monitor and identify the drone movements threatening an area to be protected (area of interest, AOI). The two clusters occupy two regions in space far apart to the right and left side of the user's view. This setting calls therefore for continuous changes in observation viewpoint to see and monitor the incoming drones. Furthermore, drones may hide each other while navigating making their recognition and tracking more challenging. Users need to monitor drone's navigation path and promptly signal hostile drones entering the AOI. Each drone follows a path chosen among a predefined set of paths. There will be only two drones among those present in both clusters that will enter the AOI. The user needs to indicate them mentioning their number (Accuracy factor). The time employed for such identification is recorded (Time factor). Figure 29 shows an example image of the operating scenario.

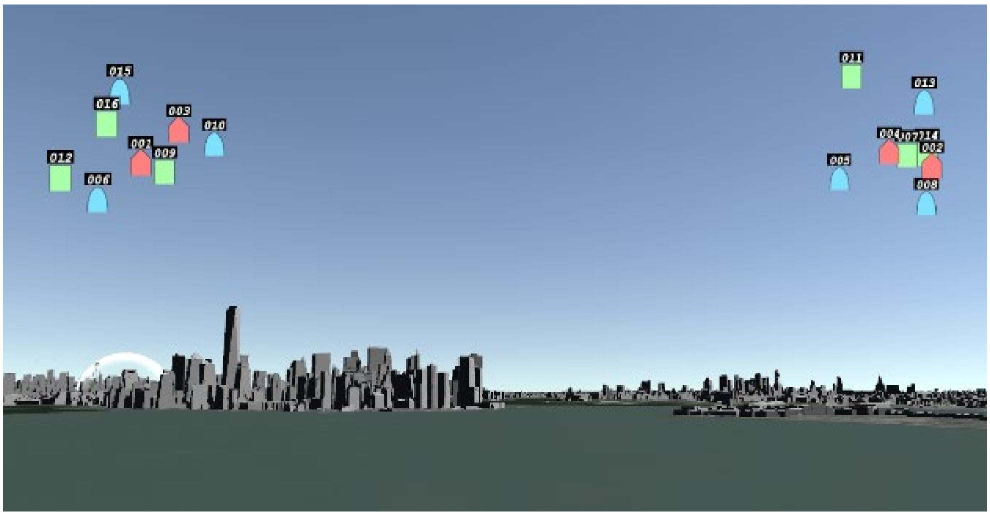


Figure 29 – Example image of the operating scenario. Colored icons represents drones. Two drone clusters are shown to the left and right side of the image.

4.2.2.3 System Apparatus

The hardware used in our experiments includes a high-spec PC with nVidia GPU GeForce RTX 3080Ti graphic card. The software uses the Unity ArcGis SDK beta 0.3 and the proprietary ELT simulator “LOKI” [75] to design and playback mission scenarios. The VR headset is an HTC Vive Pro Eye. Interacting through a joystick is through a typical console-game controller (Sony PS5 dual-controller).

4.2.2.4 Experiment System and Design

Twelve participants conducted a within-subjects' evaluation over the two systems. Participants had none to good experience with VR systems and computer games (uniformly distributed). Their age ranged between 20 and 55 (an average of 27.4 years old), and they were 53.3% females. The testing scenario and system sequence were assigned according to a predetermined counterbalanced schedule to avoid fatigue and learning effects. Our test was designed following literature recommendations and conformed to traditional approaches in terms of forms and questionnaires [49], Initially, we provided participants with information, a consent form and a pre-test screening. Next, a system practice session was administered for a user to familiarize themselves with the controls, followed by two repetitions of the operation session (R1, R2). Next, questionnaires are administered. Figure 30 summarizes our procedure.

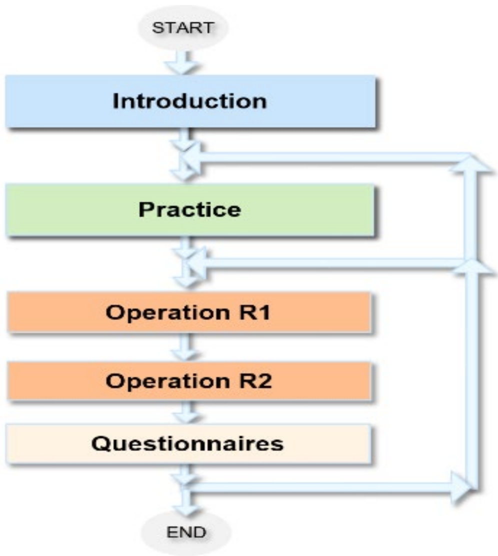


Figure 30 - Testing procedure.

We chose to rely on standard questionnaires. They were: *Presence* (Igroup Presence Questionnaire - IPQ) [50], assessing: general presence (Q1), spatial presence (Q2-Q6), involvement (Q7-Q10) and realism (Q11-Q14); *Comfort* (Simulator Sickness Questionnaire - SSQ) [51], [52]; *Ease of Use* (System Usability Questionnaire - SUS) [53]. In addition, we designed our questionnaire for *Depth Perception and Learning* addressing: general depth impression; distance judgment; ease of finding; and space objects identification inspired by [36]. Answers were provided on a 7-point Likert scale, except for the SUS (5-point) and SSQ (4-point). We computed mean values of acquired data and measured the statistical significance of results by estimating the Student's t-distribution for paired comparison with repeated measures. When considering statistics sensitivity, we set the p-value= 0.05 as the threshold. We also estimated each comparison's standard error of the mean (SE). The results are presented and analysed according to the research questions in the following sections.

4.2.3 Results and Analysis

4.2.3.1 Accuracy and Time

Table 5 shows results. The two systems show good accuracy with a higher mean value on the HMD-J and no significant differences. Concerning the time employed to discover the drone threat, the HMD-J performed significantly worse, with users needing more time (*p*=0.039). It is surely more intuitive and, therefore quicker turning the head than using a joystick, an operation the application calls for very often. Interestingly, being quicker in observation movements does not lead to higher accuracy.

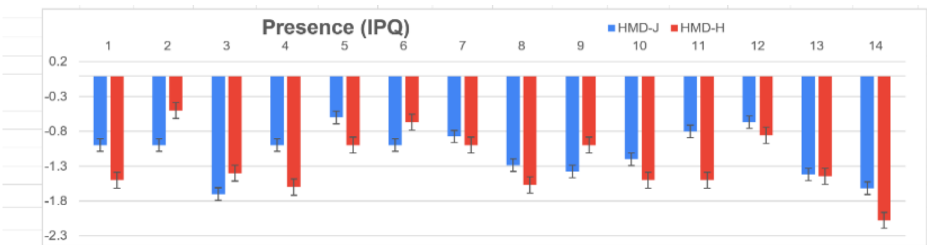
	Accuracy	Time
HMD - J	0.885	19.247
HMD - H	0.792	3.644

	Accuracy	Time
HMD-J vs HMD-H	0.356	**0.039**

Table 5 - Mean (left) and Student's p-values, (right) for Accuracy and Time.

4.2.3.2 Presence

Results in Figure 31 show contrasting indications for HMD-J and HMD-H, with the latter generally prevailing over spatial presence and realism. Furthermore, the HMD-H performs significantly better on realism (Q11 *p*=0.049). As we are testing head rotation, we would assume that the involvement subscales would be more effected than the realism. Results seem nonetheless to indicate that isolation plays a strong role on both systems and that the ability to interact with the virtual environment triggers involvement regardless of the naturalness of viewpoint change. Natural head movements still represent a more realistic way to interact, and supports this presence element. Presence questions are shown in a Table 2.



	1	2	3	4	5	6	7	8	9	10	11	12	13	14
HMD-J vs HMD-H	0.393	0.277	0.500	0.196	0.325	0.289	0.467	0.163	0.429	**0.049**	0.100	0.205	0.434	

Figure 31 – Mean, SE and Student's p-values for Presence.

4.2.3.3 Comfort

Figure 32 shows results for mean and *p*-values. Both systems have mean values between slight to moderate discomfort and almost entirely below its the mid-value. Significant differences are recorded for Q5 (*p*=0.034), Q11 (*p*=0.046) and Q14 (*p*=0.048) where the HMD-J perform worse in terms of difficulty to focus, blurred vision and vertigo (loss of orientation). This clearly appears a side-effect of joystick movements, which causes scene movements uncoupled to head movements. Comfort questions are shown in a Table 3.

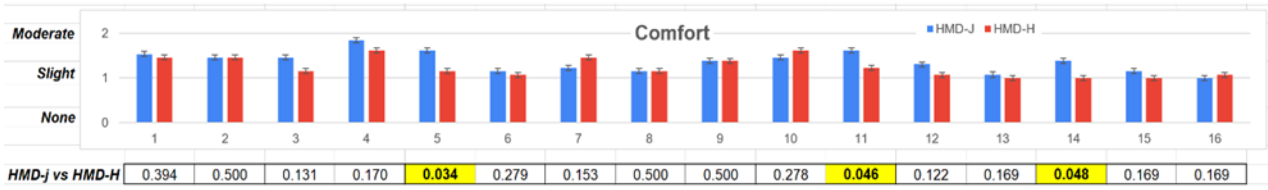


Figure 32 - Mean, SE and Student's p-values for Confort.

4.2.3.4 Ease of Use

Figure 33 shows high mean value for both systems which indicate they are liked and user-friendly. Both HMD systems also score high as well-developed product and quick to learn. The HMD-H scores are generally higher than HMD-J (except for feeling confident) and significantly more "natural to use" (Q3 p=0.035). HMD-J scores significantly higher as "awkward to use" (Q8 p=0.038) and unnecessary complex (Q2 p=0.047). Ease of Use questions are shown in a Table 4.

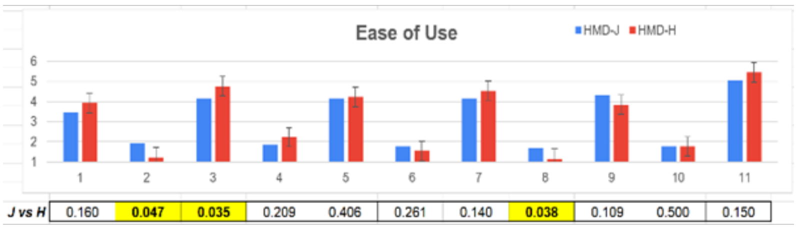


Figure 33 – Mean, SE and Student's p-values for Confort.

4.2.3.5 Depth Perception and Learning

Figure 34 shows both systems score high on all questions with no significant differences. This indicates the two control modalities do not affect depth impression and perceived learning in identification and memory recall. Depth Perception questions are shown in a Table 6.

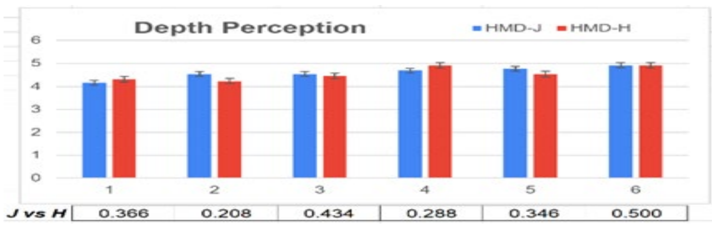


Figure 34 – Mean, SE and Student's p-values for Depth Perception

Depth Perception	From (0) to (6)
1) How much would you score the overall sense of depth impression (i.e. delivered sense of three-dimensional appearance)	Not at all -- Very much
2) How much does the tested system help with positions and distance judgement?	Not at all -- Very much
3) How much does the tested system help with spatial understandig?	Not at all -- Very much
4) How much does the tested system help with sfinding/identify/classifying symbols?	Not at all -- Very much
5) How much does the tested system help with memory and recall?	Not at all -- Very much
6) How do you bellieve the tested system help with learning/training planning or prediction?	Not at all -- Very much

Table 6 - Mean, SE and Student's p-values for Depth Perception.

4.3 Conclusions

User Study 1: Pilot-Training

This study aimed to identify and learn more about possible factors that contribute to the effectiveness of pilot-training like tasks and VR systems to understand when HMDs can be proposed. This study compared performance using DM and HMD for pilot training. Different actions were assessed for FDT and STP with different levels of difficulty.

The objective data and subjective judgements highlight similar characteristics across the research question. The outcome was quite clear: the HMD is best for more challenging training actions. It excels both in terms of precision and completion-time. The HMD performs excellently in the S&F tasks and helps the user perform faster too. HMD performance with drone teleguide is also positive in terms of precision. There is no significant difference on execution time, but it scores 52% higher (in mean value) in terms of precision of operation compared to DM. The HMD good performance is supported by the good subjective ratings it gets for presence, comfort, and ease of use.

The data also clearly show the advantage of the isolation element in terms of presence and depth impression, with the HMD performing better than DM. It is possible then to confirm the role of isolation, FOR and 3D visualization, which bring relevant advantages to HMD systems. HMD definitely supports presence and depth impression, and we have a wide set of literature works showing that presence and depth perception brings great advantages to user performance.

The DM is still very suitable for simple and repetitive actions, such as operating our check list, and for driving into a simple narrow tunnel. The DM also scores generally high and does particularly well in terms of comfort.

It seems immersive technology is ready to be widely adopted in pilot training, where its superiority against the DM becomes relevant

User Study 2: User-Interaction

This work compared the performance of two HMD systems with different interaction modalities for viewpoint change (HMD-J uses joystick, HMD-H uses head). The HMD-H being the most used, HMD-J seemed worth to be considered for applications requiring continuous and rapid change of views, as they cause great neck movements and a possible lack of situation overview.

The evaluation provided insight on several aspects:

User performance was mainly affected in terms of lapsed time. Natural head movements are intuitive, therefore faster. Rapid actions may however be critical in specific application instances, which gives HMD-H a clear advantage in terms of response time. If decision time is rather not time-constrained, HMD-J can be used, achieving the same operation accuracy. A drawback of HMD-J is being less natural to use and with some element of mild discomfort, which makes HMD-H generally preferred. Worth to note, some users experienced nausea, dizziness, and fatigue (not significant, but emerging), on the HMD-H, because of the several head movements. This did not occur with HMD-J.

The outcome for depth perception and learning aspects such as identification and memory recall, is interesting. Those aspects clearly appeared not related to the viewpoint change modality. The same can be said for accuracy of operation.

Although the HMD-J remains an unusual setup which does not seem to bring general advantages to C2 applications, the study reveals some interesting outcome which have been described.

5. The Photograph Advantage: Screen, Location and Sense of Place

There is today a new opportunity to take experience of remote places to an entirely new level thanks to latest progress in camera and display systems, and to the newly available HMDs, providing great mobility and visual performance at much lower cost. This, together with the new needs for telework and teleoperation, has brought a renewed interest towards more effective ways to enhance telepresence using VR technologies.

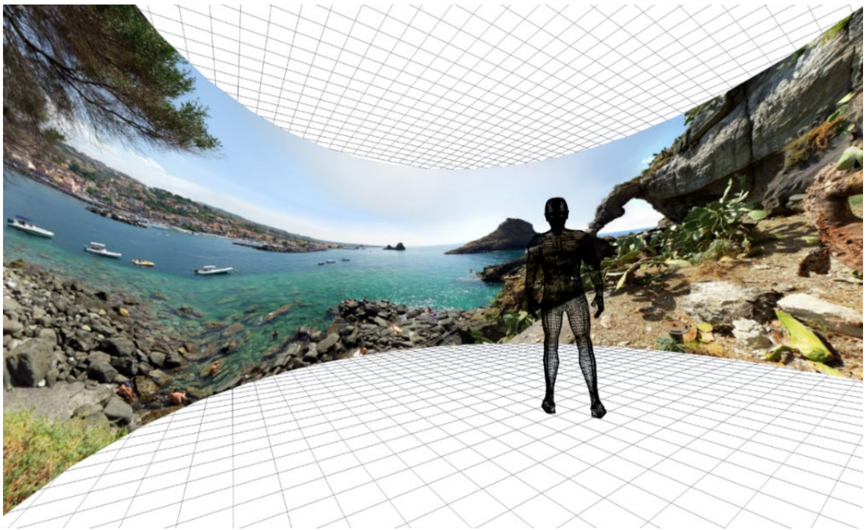


Figure 35 – Photo Base Immersive experience.

We know from the VR literature that the more immersive the telepresence experience is, the more effective a task performance is expected to be [76], e.g. in cases of psychomotor tasks [77], immersive analytics [78] and decision making [79]. However, it is not always straightforward how to maximize telepresence. It seems to be relevantly about the provided sense of presence, which in turn calls for a number of elements to contribute effectively to the provided experience.

A relevant element that plays a fundamental role in many applications, and towards presence too, is realism [80], [81]. This indicates similarity to everyday life experience. For VR systems, visual realism is of major relevance, referring to the natural viewing experience. A vivid, undistorted and correctly proportioned visual impression will enhance realism, but this objective can be hard to achieve as it calls for effective system and application design. Among main system elements that affect visual realism: image acquisition, processing algorithms, and visualization medium.

Human factors also play a role and can be affected by those system elements as well as user previous experiences [82]. We find limited literature assessing immersive observations of photographed environments and providing guidelines. A main reason being that extensive omnidirectional high-resolution capture is a very recent achievement.

The aim of this study is to investigate the contribution provided by photographic images when observing real places through a VR headset, in terms of perceived image-lighting and human sensations (visual realism, presence, emotions). The study looks at the role played by three specific elements: screen, location, and sense of place. For screen the focus is on the role played by display pixel-density on perceived image lighting by experimenting with two different screen displays. For location the focus is on the represented environments' spatial perception and on the effects of illumination by experimenting with two different locations. For sense of place the focus is on the role played by previous knowledge by experimenting with two groups of people.

5.1 State of the art

5.1.1 Visual Realism in VR

VR headsets have demonstrated being able to provide and sustain immersion and VR experience, e.g. in terms of presence and emotions, capitalizing on the portrayed image quality [83] and viewing setup [84], [85].

Literature works have focused on: (a) elements contributing to high-fidelity image reproduction and visual realism, e.g. the role played by display resolution [86] and illumination [87], and S3D viewing [88], [89]; (b) the effects of image-quality and visual realism in terms of presence, emotions, and depth perception [87] [90]. Interestingly, visual realism is also being investigated in relation to data visualization, e.g. geospatial information [91] and immersive analytics [78]. Experiments in literature are nowadays mainly performed on HMD systems. Alternatively, we find the use of 3D desktops, 3DTVs and wall screens. Visual realism is mainly assessed on synthetic images, while only few works use photographic texture.

Camera, Display, Image, Quality, Realism

Visual realism is investigated by Janssen et al. [92] and Ijsselsteijn et al. [88] in terms of image quality, which is claimed to be determined by usefulness and naturalness, a concept only partially shared by Kuijsters et al. [87], who conclude that naturalness may not contribute to image quality [93] (Seuntiens, Meesters, & Ijsselsteijn, Perceived quality of compressed stereoscopic images: Effects of JPEG coding and camera separation, 2006). Ferwerda [94] and Hagen [95] analyze the role of visual and photographic realism towards achieving a more realistic response.

Several works relate realism to different image elements. Some authors investigate S3D camera parameters most related to viewing setup. Ijsselsteijn et al. [88] focus on camera baseline, convergence and focal-length / FOV, and discusses their contribution to naturalness; whereas Banks et al. [96] focus on the appropriate matching between the camera's focal-length and image sensor-size, with display's viewing distance and screen size, and they discuss their contribution to veridical visual perception (intended as faithful representation of dimensions).

Lens and display geometry are also regarded relevant factors, as optical flow changes when display peripheral geometry is deformed by the optics, causing viewing and perspective distortions [89].

Some authors investigate camera parameters most related to *image lighting*. Kuijsters et al. [19] focus on color, contrast and texture, and discuss their contribution both to image quality and depth perception. Tiiro [97] and Pardo et al. [98] focus on color, shadows, texture and definition, to bring realism into VR scenes. Bowman et al.[86] address resolution. Slater et al. [99], [100], Palad [101], Gu et al. [102], address vivideness and sharpness.

Many authors have more recently focused on automatic objective image quality assessments with works exploiting deep learning techniques applied to images and videos, e.g. [103]. This approach, typically aimed at better streaming, can run automatically and does not require user studies. It needs parameters training, though.

Visual Realism and Presence

Visual realism often refers to different degrees of immersion and may contribute to a higher sense of presence. Literature works have discussed visual realism contribution to presence. Witmer et al. [80] discuss its role, whereas Shubert et al. [104] consider experienced realism a key element of the Igroup Presence Questionnaire (IPQ).

Interrante et al. [105] consider presence the main reason for accurate distance estimation in their photorealistic virtual environment, whereas Ijsselsteijn et al. [106] discuss presence in relation stereoscopic-3D and camera parameters. More recently, Ling et al. [107] discuss 2D and 3D viewing, FOV, center-of-projection and vantage points, and their effect over perceived presence; whereas Hvass et al. [108] discuss the effect of geometrical realism over presence.

Visual Realism and Emotions

Presence is also connected to visual realism through its effect over provoked emotions and sensations [109]. It seems that the effect of immersion over emotions may depend on the type of the emotion [83], with some authors relating immersion to arousing emotions such as fear and anxiety [84], and other demonstrating its positive effect over non-arousing emotions such as joy, relaxation, sadness and satisfaction [110].

The work of Banos et al. [110] also discuss the use of colors, reflections and natural sceneries to assess sadness, anxiety, joy, relaxation, and satisfactions, in realistically portrayed virtual environments. Seagull et al. [111], on the other hand, focus on assessing the perceived physical fidelity through quality of experience, satisfaction and enjoyability. Finally, Hakkinen et al. [112] underline how S3D image quality can affect emotions.

Visual Realism and Virtual vs Real Viewing

Some literature works focus on comparing real viewing to realistic visual replicas [113], [114], [115] and high-fidelity and faithfully-sized replicas [105], [116] and non-photorealistic replicas [117]. The work of Interrante et al. [105] discusses the effect of photorealism and realistically portrayed dimensions to distance estimation. It associates users' performance in estimating distances correctly to the perceived sense of presence and contribution of visual calibration (occurring when the correspondent real environments are also observed).

Visual Realism and Sense of Place

The user's perceived sense of place during observation is reinforced by prior knowledge of the shown environment. This has been confirmed in case of VR environments [82] and 3D geo-visualization [118]. There is a relation between emotion and sense of presence, which can alter place perception [119]. Julin et al. [120] suggest the sense of place offers an interesting point of view for assessing effectiveness of photorealistic 3D visualization. Virtanen et al. [121] find that 3D geo-visualization helps comprehension of spatial relations.

5.2 Proposed System Concept

The research focus is on delivering a realistic *true-dimensional* visualization. We define it a viewing experience that let users perceive the actual space dimension, objects' size and distance. This is different to what we typically experience when looking at environment photographs or a show on TV. In these cases, wide-angle views are often used to include large space-portions on one single view. The perceived environment and object sizes become then typically magnified, which causes surprise or disappointment when one sees the portrayed place through own eyes (direct viewing).

High fidelity visualization is a difficult process that often requires careful setup for capturing, modelling and viewing [122]. If we look at the processes taking place when visually observing a real place through a medium, we find several works trying to identify the main involved actors, e.g. [23], [123], [96]. We summarized those actors as: scene content, camera system, image processing, visual display and human sensory system.

Assuming the human sensory system is the same for all observers, and having those observers looking at the same scene, we limit our focus on the remaining actors: camera, processing and display. We wish then to look at how those actors can deliver the impression of representing the observed environment features faithfully.

The general concept for the approach is to acquire photographs or videos of a remote place and play them back appropriately to maximize realistic impression.

The referred system capitalizes on the use of high-resolution, omnidirectional and three-dimensional images. It relies on acquiring photographs or videos of a remote place and play them back appropriately to maximize realistic impression. The latter is proposed to be achieved from a system concept point of view through providing:

- *High-fidelity Images.* The visualized images should be of high photographic quality, with minimum visual distortion and faithfully reproduction of lighting and reflections, (including brightness, contrast and color reproduction). This calls for high-resolution images and displays, quality lens, and lossless processing algorithms.
- *Natural FOV*. The perceived field-of-view (FOV), key in estimating distances [23], [123], should resemble the one we naturally perceive. Typically, this does not happen in VR and distances are often underestimated [23], [124], [125], [126], including when

real visual sources are used [127], [128]. This for example happens when observing photographs captured with wide FOV, a setting that is often adopted for practical needs, e.g. when an entire room needs to be visible within a single image. Furthermore, being the structure of a camera image different from the one of human eye-retina, we are not replicating the same image formation. Nonetheless, capturing with appropriate camera focal-length helps perceiving a close-to-natural viewing angle. Also, observing through a natural FOV typically means observing a portion of the presented scene, which in turn means one can have higher-resolution (with the same camera) over the visible space portion. When observing through a VR headset we also need to consider alignment between application-camera FOV and display FOV. The latter should be wide enough to cover for eye's peripheral vision. This result can be achieved with relatively large displays or by lens zoom-in, leading to have invisible screen's edges. One should then be careful this may cause a portion of display area to be hidden. To avoid the latter, a careful FOV adjustment is needed for large displays, which in turns means being able to measure display's visible portion, and then set the application-camera FOV accordingly [113]. In summary, natural FOV involves matching software, headset and human FOVs.

- *3D Vision*. The visualized images should provide binocular three-dimensional vision as it happens with direct viewing. Works in the literature have demonstrated the advantages of S3D compared to 2D viewing over several factors, e.g. naturalness [87], [88], depth realism [129] [78], spatial perception [130], and improved performance [36], [85]. The acquired stereoscopic couples need to be captured with a baseline (distance between camera's objectives) close to human inter-ocular distance (IOD). Furthermore, cameras should possibly be set to parallel configuration to reduce distortions and increase comfort [88], and there should be close match between focal length/sensor size and viewing distance/display size [96]. The contribution of 3D vision to visual realism and naturalness is supported by many literature works [87] [90].
- *View by Head-Rotation*. The panoramic-view has to be seen/discovered by rotating our heads. This represents the way we naturally see the world. A VR headset allows for the same viewing gesture. What we need is to display coherent neighbor images in continuous fashion. This needs to be carefully designed (and engineered). Viewing by head-rotation can greatly contribute to VR display fidelity, realism and spatial understanding [86].

In summary, VR headsets can allow for panoramic view observation by head rotation, while keeping natural FOV, and they can also allow for high-resolution images displayed three-dimensionally (through separate viewing displays). Clearly, achieving natural viewing behavior also depends on natural image response to head rotation. The performance of headset's inertial, magnetic and visual sensors, is therefore also relevant. If we wish true-dimensional visualization to be achieved, we also need to focus on lens characteristics (and distortions). This helps towards realistic perception of distance to objects (egocentric) and among objects (relative).

5.3 Investigation and Research Questions

An investigation on VR portrayed photo-based realistic viewing, therefore involving both image acquisition, processing and visualization, may easily lead to complex and extended investigations due to the many system parameters involved. These include those related to the type of optical system the cameras and viewers have, acquisition and visualization setups, software processing, graphical rendering, image corrections and view-porting algorithms, etc. Clearly, an exhaustive investigation is realistically impossible, and we need to largely delimit the variable parameters based on available resources (funds, and above all, time).

In addition, there is a degree of difficulty in isolating specific system elements of VR systems, and this is even more difficult in our case as we target practical applications, and therefore wish/need to use off-the-shelf ("inflexible") hardware.

A major question is then about whether we can still produce meaningful results with a delimited resource instance. This question has definitely been asked and the answer has been that one can still have meaningful results by both relying on same literature outcome and by designing the experiment such that it would constraint possible outcomes. The literature indicates:

- *Concerning screen displays*: a major role is played by display viewing-setup, image-definition and lighting.
- *Concerning location environments*: a major role is played by the portrayed range of light-intensity/colour, and distance to objects.

Experimenting with different VR headsets (as the HTC and Oculus) would cause many parameters to change among which those related to viewing-setup. This is the reason the same VR goggle system has been proposed, (so that lens, eye-display distance and visible surface, were identical), so as the smartphone-based system (so that we could have different pixel-density and lighting quality). In addition, the display brightness has been controlled and made it equal on both displays using a LUX meter. This way, only display pixel-density and illumination characteristics would represent the main difference.

Concerning the experiment design for system display, the display choice was made such that the system with lower pixel-density would have better lighting characteristics. The iPhone display was then chosen. This was done to ensure that a possible lower performance on this display was generally not to be attributed to display lighting characteristics. Analogously, a better performance from the display with higher pixel-density (LG display) was not to be attributed to better display lighting characteristics.

As for the environment, 2 environments have been chosen such that would clearly have different ranges of light-intensity and distance to objects. Minding our target application, the proposed island and cave views have been chosen. In particular, the Island has a high range of colours and distances, whereas the Cave has a low range of colours and distances. This was done to ensure that if the island view would have induced more realism, this would have mainly to be attributed to the wide-range light-intensity. Analogously, a more realistic performance with the cave view would have mainly to be attributed to the closer distances (rather than its nearly monochromatic views).

The above-described settings are coupled to our testing strategy, which relies on pairwise comparisons on either the same display or the same environment. This way the possible outcome is constrained, and results are exposed in a clear manner.

We decide then to investigate the effect of some display and environment elements to realistic viewing. We focus on: display pixel-density, and environment illumination and distance to objects. We add to our investigation a specific study looking at the role of previous knowledge, based on what we have learnt on sense of place from literature works.

Research Questions

The research question involves the following aspect:

- *Screen*. How display pixel-density affects realistic viewing experience?

  The literature has shown that display quality is connected to its lighting characteristics (color, contrast, etc.), and these affect some human factors (realisms, presence, etc.). We focus on the effect of display pixel-density towards realistic viewing.

- *Location*. How the perceived sense of realism varies for different types of locations environments?

  The literature has shown that the portrayed environment characteristics affect some human factors. We focus on the effect of environment illumination and spatial perception.

- *Sense of Place*. How place familiarity affects realistic viewing experience?

  The literature has shown that previous knowledge may reinforce or alter perceived presence. We focus on its effect to visual realism.

The aim is to answer the above research questions (representing independent variables) by asking users to judge on number of factors (dependent variables) that can potentially affect them.

Subjective ratings and measurements represent trustable indications for immersive VR systems due to the high user involvement and observable effects those systems provide [105].

The ratings are related to display perceived lighting and human sensations in terms of visual realism, presence and emotions. The measurements are related to the perceived distance to objects.

The proposed study is unprecedent and merely explorative. Previous research on contribution of visual realism when observing omnidirectional three-dimensional photographs of real places through a VR headset, is hard to be found. We hope the study can raise awareness on the effect that photo-based VR representations have towards visual realism.

5.4 Evaluation Measurements and Experiment Design

5.4.1 Display Factors

The difficulty in isolating a specific system element is typical when assessing VR systems. This is even more difficult nowadays as we predominately use off-the-shelf systems. There is awareness this represents a limitation in our study but believe we can still provide meaningful results as seen in many literatures works. Bowman et al. [86] discuss this issue for image-resolution in support of multiple-component assessments.

In this study display quality is assessed by interrogating users on a number of *image lighting* characteristics. We choose the below listed elements, and for each element we ask users to rate the element’s perceived *relevance to realistic viewing*.

- *Pixel-Density*. Overall pixels’ number divided by screen-size. It is related to image resolution and provides an indication of the perceived degree of detail. It affects image illumination too.
- *Lightness*. Overall perceived light intensity. It typically depends on display luminance (amount of light radiance) and the set display brightness.
- *Color*. Combination of light’s hue and intensity. It represents the wide variety of color shades.
- *Contrast*. Difference in lightness between pixels. It typically indicates the overall tonal range, which distinguishes between brightness and darkness. The perceived contrast in photographs depends on the location of where shadows and highlights occur [101].
- *Vividness*. Clarity, richness and liveness of the image. It typically represents the contrast of image mid-tones (leaving shadows and highlights unchanged) enhancing the appearance of overall details.
- *Sharpness*. Distinctiveness among pixels. It indicates how well pixel borders merge together and therefore the perceived focus (level of detail).
- *Definition*. Absence of blurs and pixelation. It typically indicates clarity of all represented objects.

5.4.2 Human Factors

Human Factors are relevant because of their relation to visual impression and to display factors. Users are interrogated on the perceived: visual realism, presence and emotions. Users are also asked to estimate distance to objects to assess its potential role in realistic viewing experience.

Visual Realism

There is no standard approach to determine visual realism in VR applications [131] [132], but few inspiring examples mostly related to specific applications. A number of authors propose the use of image-quality metrics and subjective rating to assess visual realism, image quality and naturalness. Brackney et al. [133] indicate interaction, control and motion, as relevant elements that should be part of a realism questionnaire and proposes the use of the INASCL simulator [134], which contemplates 11 criteria including fidelity, defined as true-to-life experience. It divides realism as physical, conceptual and psychological, with all those aspects to contribute to engagement and to be included when assessing realism of VR simulations. Wilson [131] proposes the Realism Assessment Questionnaire, partially inspired by Hill [135], and applies it to colonoscopy studies.

It is proposed to interrogate users on four questions, taken from the *experienced realism* subscale of the Igroup Presence Questionnaire (IPQ), also referred as *realness* [104], [136]. Two questions are then added related to photorealism. Questions are summarized below:

- *Realness*. (a) Similarity to real world; (b) experience similar to real world; (c) similarity to imagined world; (d) excessive realism [136].
- *Photorealism*. Overall level of photorealism.
- *Similarity to photo/video*. Experience similar to seeing a photo or video.

Presence

Different methods have been proposed to assess the sense of presence, which are typically based on subjective ratings given through validated questionnaire. Among the most popular, Slater, Usoh and Steed [99], Witmer and Singer [80] and the above mentioned IPQ [136]. It is proposed to rely on the use of the IPQ.

Therefore, the questions address the followings:

- *Overall Presence*. (P1) Sense of being there.
- *Spatial Presence*. (SP1) Sense of surrounding reality; (SP2) perceiving pictures only; (SP3) feeling present in virtual space; (SP4) sense of acting there; (SP5) feeling present in virtual environment.
- *Involvement*. (INV1) real world awareness; (INV2) real world unawareness; (INV3) attention to surrounding reality; (INV4) attention to VR world.

Emotions

Authors have typically investigated this aspect through questionnaires targeting specific emotions and visual elements expected to elicit emotions, e.g. image characteristic such as light intensity, shadows and colors. We propose a number of potentially relevant emotions (selected after a pilot assessment) and ask users about their *current sensation towards the emotion*. At the end of the entire evaluation, we also ask users about their: *sensation of being back to reality*. The selected emotions are listed below.

Emotions. Happiness; Enjoyment; Relaxation; Scariness; Sadness; Anxiety; Anger; Surprise; and Back to reality sensation.

5.4.3 Evaluation Design: User Study 3 and User Study 4

5.4.3.1 User Study 3: Screen and Location

A within-subject study was conducted with 20 test-users. Participants' age ranged between 22 and 53, with an average of 28. The study had the conditions described below.

- *Screen*. Two displays were chosen that mainly differed in pixel-density, aiming to gain insight on how much this element can affect the considered display and human factors. More details about displays and their choice are in section 4.3.
  - *LG display.* It has higher pixel-density.
  - *iPhone display*. It has lower pixel-density.
- *Location*. Two environments were chosen which mainly differed in illumination conditions and distance to objects. The aim was to gain insight on how much those characteristics can affect the considered display and human factors.
  - *Island*. It portrays a wide range of light-intensities and colors. It has a high range of distances to objects.
  - *Cave*. It portrays a limited range of light-intensities and colors. It has a low range of distances to objects.

All combinations of the above conditions related to display and environment were tested by running pairwise comparisons on displays (A1 vs B1, A2 vs B2, A1+A2 vs B1+B2) and environments (A1 vs A2, B1 vs B2, A1+B1 vs A2+B2).

5.4.3.2 User Study 4: Sense of Place

An in-between subject study was conducted with 40 test-users. Participants' age ranged between 25 and 58, with an average of 32. Participant had none to good experience with VR devices and computer games (uniformly distributed). The study had the conditions described below.

- *Familiarity*. We chose two user groups that differed for their prior knowledge of the shown environments, aiming to gain insight on how much familiarity with the observed place can affect visual realism and presence. The study had the conditions described below.
  - *Site-Familiar*. Users that knew well the observed site.
  - *Unfamiliar*. Users that had never been in the observed site.

Two above conditions above were tested by running pairwise comparisons on the two environments. We tested on one display only for practical reasons. We had therefore the following combinations: C1 vs D1, C2 vs D2, C1+C2 vs D1+D2.

The site-familiar users were people that worked in the considered environment and visited the place on a weekly or monthly base. They were therefore ideal to assess faithful reproduction of environment objects as they knew them well. We thought their given scores would be particularly meaningful (a sort of “ground truth”). Unfamiliar users could still judge realism based on appearance of well-known objects such as trees and landscapes. We considered of interest to see how site-familiar users’ scores would differ from those of unfamiliar users.

5.4.4 Test Procedure

The test organization and procedure followed literature recommendations [137], [49]. Users executed under the same conditions the tasks described below. The Figure 36 shows an illustration of the main experimentation steps. Table 7 shows what factors were assessed in each study.

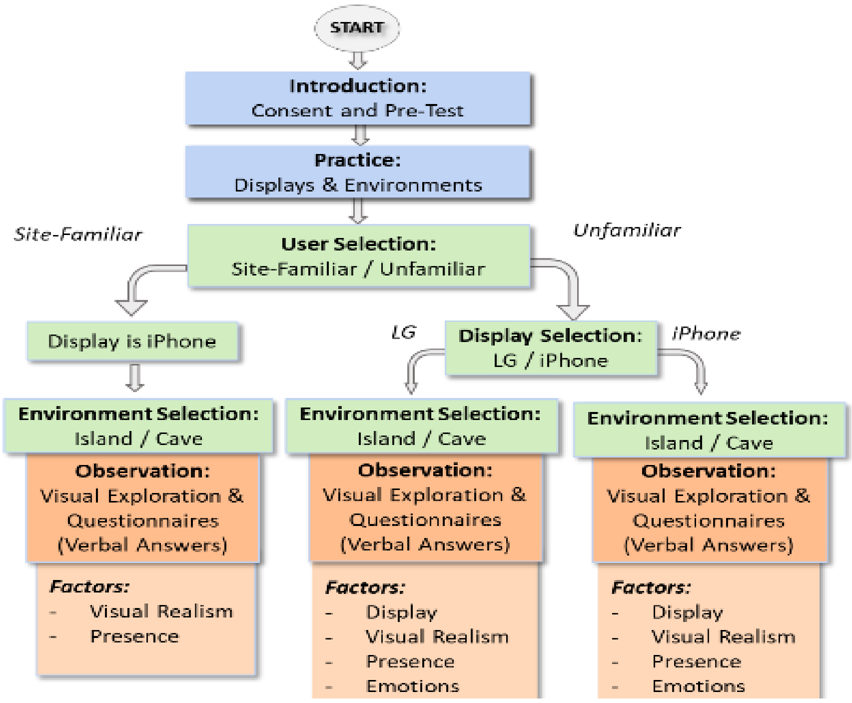


Figure 36 – Main experimentation steps.

Display and Human Factors	Display	Environment	Familiarity
Image lighting	X	X	
Visual Realism	X	X	X
Presence	X	X	X
Emotions	X	X	

Table 7 - Display and human factors assessed in the experiments.

- *Introduction*. Initially, we provided participants with an information sheet, consent form and pre-test screening (background questionnaire and a vision test using Snellen chart to check if any sight issue. We carefully explained participants about the meaning of each evaluation factor and the related questions. As image-lighting characteristics have clear visual meaning, we supported questions with illustrations.

- *Practice*. Participants were asked to familiarize with tasks and system through practice trials. Participants were asked to explore panoramic views while wearing the VR headset. Each participant ran trials on each display and on each environment.
- *Selection*. The task and environment sequence were assigned to each participant according to a pre-determined schedule (based on the Latin square root) to counterbalance task sequence and avoid fatigue and learning effects.
- *Observation*. Participants correctly, firmly, and comfortably, wore the headset. They were then observing the panoramic view for as long as they needed. During the observation participants were questioned according to the set factors and provided answers verbally.
- *Questionnaires*. We conformed to the traditional approaches in terms of forms and questionnaires [49], [138], [139]. Questionnaires were designed according to the 7-point Likert's scale and scores range was -3 to 3. Forms included sections for reporting open comments through written feedback.
- *Test Results*. We computed scores' median and standard error. We also measure statistical significance by estimating paired Student's T-test. There is some debate on whether Likert scale variables can be treated as ordinal or categorical, nonetheless some piece of research choose to see it as a continuous variable. In our studies (with 20 users in a within-subject setting and 40 users in between-subject setting) we verified a normal distribution of data, and we then computed the p-value according to the Student's T test [140]. We measured the effect of the different displays, environments, and familiarity on the dependent variables related to display and human factors. We sat that an alpha of 0.05 as p value, determined whether the result is judged statistically significant (tables' red numbers). Alpha values between 0.05 and 0.06 were referred as having a "tendency to significant" (tables' brown numbers). The results are presented and analyzed below according to each research question and related pairwise comparisons. Evaluation measurements are mentioned to perform: high for median score between 1.5 to 3; low for median score -3 to -1.5; and medium for median score -1 through 1

5.4.5 Experiment Setup and Apparatus

Literature works have indicated that naturalness and realistic space perception are relevant for presence [92], [94], [105], [106], therefore for the system used in this research a careful selection of off-the-shelf hardware has been made. The identified choices were confirmed by running a number of pre-assessments.

5.4.5.1 Acquisition System

The choice for the image capture-visualization system was to have a technology that would feature high-resolution, omnidirectional, and correctly proportioned three-dimensional images [141].

For the image acquisition we chose the consumer 3D camera *Fujifilm 3D FinePix W3* because, among those available in the market, it provides focal-length and camera stereo-objectives separation (baseline) closer to the average human inter-ocular distance IOD (6.5 cm). A dense set of images were needed to keep distortion low. Therefore, hundreds high-resolution images of the environment were captured at set locations by rotating the camera of approximately one degree, up to cover a 360-degree view. Single overlapping images merged through a stitching process into one high-resolution panoramic composition using our algorithm based on SIFT-like feature matching [18]. The obtained panoramas were organized after spherical configuration from single viewpoint, with stereo-couples going through the same stitching process to avoid mismatches that can affect 3D viewing comfort.

5.4.5.2 Visualization System

The smartphone-based VR goggles was the preferred choice because it allowed us to test with different displays, while keeping the same remaining headset's characteristics, such as lens and eye-display distance. We chose the VR Shinecon glasses. Our headset allowed users to adjust focus and IOD. We chose the LG G-series and the Apple iPhone (the display characteristics are in Table 8). Both displays were IPS LCD.

The two displays differed in size (the LG was 17% larger). However, the additional display surface was not visible to users because a display portion was hidden by the headset structure. This resulted in users observing the same FOV on both displays (we measured a FOV discrepancy below 5%). The LG display had greater pixel-density compared to the iPhone, whereas the iPhone display excelled in lighting characteristics such as: black levels, contrast, grey scale, saturation and color accuracy. We chose the system with lower pixel-density (iPhone) to have better lighting characteristics to ensure that a possible lower performance was generally not to be attributed to display lighting characteristics. Analogously, a better LG performance was not to be attributed to better display lighting characteristics. We made sure display brightness was set as equal as possible using a LUX meter. Figure 37 shows the VR headset. The Table 8 shows the display specifications.

Display Characteristics	iPhone	LG G-series
Max brightness in Nits (Higher is better)	**559**	379
Black levels in Nits (Lower is better)	**0.3647**	0.4337
Contrast Ratio 100% Brightness (Higher is better)	**1,534.0**	875.0
Avg. White point in K (Closer to 6504K is better)	**6,515**	7,244
Greyscale Accuracy (Lower is better)	**1.9683**	3.6935
Saturation Accuracy (Lower is better)	**1.19.29**	4.7599
Great MacBeth Color Accuracy (Lower is better)	**1.7645**	3.9702
Display Resolution (pixels) / Screen Size (in.)	1334x750/4.7	**2560x1440/5.5**
Pixel Density (ppi) / Pixel Size (mm)	326/0.078	**538/0.047**

Table 8 - Technical specs of our displays (better values in bold).

5.4.5.3 Location Environment

Two location environments were chosen with different illumination conditions. One depicted a wealth of light-intensities with wide-color and distance ranges. The other one had opposite characteristics, i.e., low ranges of light-intensities, colors and distances. The reasons for this choice were in the fact that the works in literature show that colors, as well as closer 3D views, are expected to increase visual realism, presence, and spatial perception.

Therefore, if the first environment would have induced for example more realism, this should be mainly to be attributed to light intensities and colors. Analogously, a more realistic performance in the second environment would be mainly to be attributed to the closer distances (rather than the nearly monochromatic views).

The chosen locations were two nature reserves located in Sicily (Italy). They were the Lachea Island [142] (further referred as island), and the Monello Cave [143] (further referred as cave). Figure 37 shows images of the two chosen environments [144].

Figure 37 - The two environments observed in our experiments. The Lachea island Invalid source specified. (top and middle left rows) representing richness of light-intensities with wide-color range and higher distance-range. The Monello cave (bottom rows) representing low light-intensity and color ranges, and lower distance-range. The used VR headset is shown in the middle-right row.

5.5 Results and Analysis: Screen Pixel-Density

This section analyses results focusing on the role played by the screen display. Scores were gathered from users looking at the same location environment (either the island or the cave) through different displays (LG and iPhone). Referring to what is described in section 4.1 this means: A1 vs B1; A2 vs B2; and A1+A2 vs B1+B2. Figure 38 and Tables 9 through 12 show the obtained results.

Figure 38 – Outcome for User Study 3 showing median values (with standard error): Image-Lighting (top-row); Visual Realism (2nd row from top); Presence (3rd row from bottom); Emotion (bottom-row). The displays are indicated as LG and iPhone, the environments as Island and Cave. For Display-related pairwise comparisons, pay attention to (A1 vs B1), (A2 vs B2) and (A1+A2 vs B1+B2). For Environment-related comparisons, pay attention to (A1 vs A2), (B1 vs B2) and (A1+B1 vs A2+B2)

5.5.1 Display Factors

Results: Relevance to realistic viewing scored high on both displays on image contrast, sharpness and definition; and on the LG display only on image lightness and color. All other image characteristics scored medium. The LG display performed significantly better than iPhone the perceived relevance to realistic viewing provided by: pixel-density (scores of both environments combined led to p=0.0482, whereas scores from the island only led to p=0.0306); color (scores in the cave led to p=0.043, and scores in both environments led to a tendency to significant, resulting in p=0.059; vividness (scores in the island led to p=0.0248); sharpness (scores in the cave led to p=0.0473); and lightness (scores in the island led to a tendency to significant, resulting in p=0.0538). Table 9 shows all p-values.

	Pixel-Density	Lightness	Color	Contrast	Vividness	Sharpeness	Definition
Isl. & Cave	**0.0482**	0.1068	**0.0594**	0.2824	0.2624	0.2736	0.3688
Island	**0.0306**	**0.0538**	0.0759	0.3658	**0.0248**	0.5000	0.2951
Cave	0.0658	0.1598	**0.0430**	0.1989	0.5000	**0.0473**	0.4425

Table 9 – Image lighting: LG display vs iPhone display (p values).

	Similarity to Real World	Exp. Similar Real World	Similarity Imag. World	Excessive Realims	Photorealism	Similarity to Photo/Video
Isl.&Cave	**0.0514**	0.1537	0.1566	0.3377	0.2601	0.3934
Island	**0.0539**	0.2718	**0.0045**	0.2663	0.4310	0.2869
Cave	**0.0049**	**0.0357**	0.3086	0.4091	0.0892	0.5000

Table 10 - Visual Realism: LG display vs iPhone display (p values).

Analysis: The high scores on a few factors and the medium scores on the remaining one, indicate that users were generally satisfied with both the displays and considered them of high quality. The scores on pixel-density indicates the higher LG display's pixel-density was noted by the users, and has therefore potential to play a role towards other factors. As for color performing higher on LG display, this is contrary to what one would expect looking at display related specs. The above results and some literature works stating display resolution and pixel-density enabling for more detailed reproduction of color shades and contrast [86], [87], [145], make us think it is the LG higher pixel-density causing higher perception of color contribution to realistic viewing. The higher pixel-density appears therefore to outweigh the better iPhone's display color accuracy and higher specs in terms of: grayscale, saturation, black and white levels, and contrast ratio. As for vividness, lightness and sharpness, there is clear indication of their contribution to perceived realism.

Overall, the outcome of display factors represents a new finding not specifically addressed in the literature. It shows a dominant role of display pixel-density compared to the display lighting characteristics.

5.5.2 Human Factors

5.5.2.1 Visual Realism

Results: The scores of similarity to real world and experience similar to real world were all high, with top values on LG in the cave (median 3 and 2.5 respectively). The first factor recorded a significant better performance of the LG display in the cave (p=0.0049) and a tendency to significantly better performance in the island and on both environments (p=0.0539 and p=0.0514 respectively). The second factor recorded a significant better performance of the LG display in the island (p=0.0045). The scores for similarity to the imagined world were medium to high with the LG display performing significantly better than iPhone in the cave (p=0.0357). Excessive realism scored low (median -3 to -2.5) and similarly on both displays. Photorealism scored high on both displays, whereas scores for similarity to photo/video were slightly lower. There were no significant differences on both factors. Table 11 shows the p-values.

	Being There	Surr. Reality	Pictures Only	Pres. Vir. Space	Acting There	Pres. in VE
Isl.&Cave	0.1270	0.0775	0.0977	0.0959	0.0711	0.1695
Island	0.1872	**0.0211**	0.1435	**0.0458**	0.0978	**0.0193**
Cave	0.0668	0.1340	**0.0520**	0.1459	**0.0444**	0.3196

Table 11 – Overall and Spatial Presence: LG vs iPhone displays (p values).

	Real World Aware	Real World Unaware	Att. Reality	Att. Vir. World
Isl. &Cave	0.3980	0.3417	0.1947	0.4608
Island	0.2960	0.2994	0.2162	0.5000
Cave	0.5000	0.3840	0.1731	0.4217

Table 12 – Involvement: LG display vs iPhone display (p values).

Analysis: *Realness* scored high-level and consistently across its four factors. Users commented of a remarkable similarity to reality, which was only undermined by the lack of environment dynamics and the limited movement options beside head-rotation. *Photorealism* was highly appreciated and often commented as “very impressive”, “effective” and “definitely greater than that provided by a photograph”. The latter explains the slightly lower scores of the *similarity to photo/video*. The overall outcome for *visual realism* confirmed effectiveness of our system, which indicates that image acquisition and processing did a good job in maintaining high quality and well combined with visualization. From users’ comments and scores, we can state that the vivid and highly photorealistic images do not replace for the lack of dynamics and missing user’s actions, but images still induce a very realistic visual experience.

Many users pointed that a correct depth impression was a key supporting element, which is in line with [88], [92]. The higher visual realism of the display with higher pixel-density (LG) is generally expected, but we have now shown this has taken place also against better display lighting specs.

5.5.2.2 Presence

Results: Both displays scored high the *overall presence* (P1), *surrounding reality* (SP1), *feeling presence in virtual space* (SP3) and *feeling present in virtual environment* (SP5). The SP1, SP3 and SP5 saw a significantly better performance of the LG display over the iPhone in the island only (p=0.0211, p=0.0458 and p=0.0193 respectively). Low were the scores of *perceiving pictures only* (SP2), with the LG display scoring a tendency to significant better performance in the cave (p=0.052); whereas the *sense of acting there* (SP4) scored medium, with the LG display scoring significantly better in the cave (p=0.044).

The *real world awareness* and *unawareness,* and *attention to the VR world* (INV1, INV2 and INV4 respectively) scored high (median 3 for INV1 and INV2); whereas the *attention to surrounding reality* (INV3) scored low (median -3). The scores for both displays were very similar, therefore no significant differences were recorded. Tables 11 and 12 show all *p*-values.

Analysis: The outcome clearly shows both displays gave users a strong sense of *overall* and *spatial presence*. The latter being undermined only by the static nature of the images, which affected the *acting there* (SP4) factor. Fourteen out of twenty users (70%) positively commented about VR headsets' suitability to scene exploration. In particular, they found natural with both displays to explore the scenes through head rotation and believed this highly contributed towards increasing spatial presence in the remote place. This is in line with the literature [80], [146], [147]. Twelve users (60%) were also particularly appreciative of the headsets' light-weight, portability and the provided sense of isolation from the surrounding environment, which they commented as also contributing to presence [100].

The significant better performance of *surrounding reality* (SP1) on the LG display only, was credited to the higher pixel-density, which is in line with some literature [107]. The low scores given to the *perceiving pictures only* factor (SP2) indicates effectiveness of presence and of visual realism too, (pictures as such were no longer noted). Despite scoring medium the *sense of acting there* (SP4), users commented the observation as "very engaging", "rich of visible elements" and "showing great variation over different viewing directions". Interestingly the SP4 significant higher scores in the cave environment of the LG display compared to the iPhone, were often commented as caused by the closer distances calling for frequent head rotations.

Regarding *real world awareness* and *unawareness* (INV1 and INV2), scores show that users clearly forgot about their actual premises once they wore the VR headset. HMDs are well-known for their isolation from surrounding space, which in our case appeared further enhanced by strong depth impression (leading to high spatial presence). The two displays scoring high and nearly identical demonstrate good quality. The *attention to surrounding reality* (INV3) and *attention to virtual world* (INV4) scoring opposite (low and high) are what is expected to demonstrate great involvement.

In summary, both displays seemed to be able to convey high presence despite the static content, with the LG display in most cases performing better than the iPhone in bringing a greater surrounding sensation (in the island) and a stronger sense of acting (in the cave).

5.5.2.3 Emotions

Results: Users scored high on both displays *happiness*, *enjoyment* and *surprise*, and on LG display *relaxation*. Users scored medium all the other factors except for *anger*, which scored low. There were no significant differences between the displays, except for *anxiety* in the island where the LG display scored significantly higher (p=0.045). Table 13 shows all *p*-values.

Analysis: The above outcome indicates emotions are triggered and most of them are positive, which goes perfectly along with what assessed in [111] regarding perceived physical fidelity through quality of experience, and in [112] regarding S3D image quality affecting emotions. As for the significantly greater anxiety when observing the island through the LG, this may be due to the greater pixel-density. The outcome for back to reality sensation after the VR experience, was varied. To some, it induced excitement, to others disappointment. This seemed therefore to be a subjective aspect.

5.6 Results and Analysis: Location Environment

This section analyses results focusing on the role played by the shown environment. Scores were gathered from users looking through the same display (either LG or iPhone), at two different environments (the island and the cave). Referring to what is described in section 4.1 this means: A1 vs A2; B1 vs B2; and A1+B1 vs A2+B2. Tables 10 through 13 and Tables 14 through 17 show the obtained results.

	Happiness	Enjoyment	Relaxation	Scariness	Sadness	Anxiety	Anger	Surprise	Back Reality
Isl.&Cave	0.3955	0.2386	0.1289	0.2388	0.1511	0.1331	0.3784	0.1440	0.4368
Island	0.5000	0.3397	0.1546	0.2330	0.1008	**0.0450**	0.3233	0.1490	0.4206
Cave	0.2911	0.1375	0.1032	0.2447	0.2013	0.2211	0.4334	0.1390	0.4530

Table 13 – Emotions: LG display vs iPhone display (p values).

5.6.1 Display Factors

Results: The contribution of definition to realistic viewing scored high in both environments. The contribution of lightness, color, contrast and vividness to realistic viewing scored high in the island only. The contribution of sharpness to realistic viewing scored high in the cave only. As for contribution of pixel-density to realistic viewing, it scored high in the island, but only when observed through the LG display.

We can observe that scores in the island were significantly higher compared to the cave: on both displays in terms of contribution to realism given by pixel-density, lightness and color (respectively p=0.0058, p=0.0212, p=0.0046); and on each single display, (except for color on the LG display where only a tendency to significant difference was recorded, p=0.0567). There was no significant difference for contribution of lightness on the iPhone. The contribution to realistic viewing of vividness scored significantly higher in the island (p=0.0401), whereas the contribution of sharpness scored significantly higher in the cave on the LG display and on both displays (equally with p=0.002). Table 14 shows all p-values.

	Pixel-Density	Lightness	Color	Contrast	Vividness	Sharpeness	Definition
LG & iPhone	**0.0058**	**0.0212**	**0.0046**	0.4641	0.1507	**0.0020**	0.0680
LG	**0.0308**	**0.0286**	**0.0567**	0.3113	**0.0401**	**0.0020**	0.2145
iPhone	**0.0361**	0.1484	**0.0146**	0.2608	0.3960	0.1100	0.0941

Table 14 - Image lighting: Island vs Cave (p values).

Analysis: If we compare data from the two environments on both displays combined (first row in Table 8) with data from each individual display (so either LG or iPhone, respectively second and third rows), we note the significant differences occurring when combining data of both displays still occur on each individual display only in case of pixel-density and color. This shows the environment plays a major role compared to display on those two factors.

According to users' comments, the contribution of pixel-density to realistic viewing was very appreciated in the island because objects were looking well defined even at the far distances this environment portrayed. The above indicated contribution of pixel-density is related to represented object distances. As for color contribution to perceived realism, it was very appreciated in the island because of the wider color-range and warmer tones. This indicates scene lighting plays a role towards contribution of color to perceived realism.

As for contribution of lightness and sharpness, the significant differences between environments when data from both displays are combined, were confirmed on LG only. The contribution of vividness was significantly different on the LG display only. The above outcome tells us the role played by the environment over lightness, sharpness and vividness, is subject to the role played by the display. In other words, we can say that we need the LG display higher pixel-density to trigger a significant difference in the contribution of those factors to perceived realism.

5.6.2 Human Factors

5.6.2.1 Visual Realism

Results: The two environments showed high median values over all visual realism factors except for the excessive realism which scored low. There were no significant differences in all the realness factors except for the similarity to the imagined world, where the contribution to realism in the island scored significantly higher when observed through the LG display only. The contribution of Photorealism scored significantly higher in the cave on both display and on the LG display only (p=0.0369 and p=0.0397 respectively). Table 15 shows all p-values.

	Similarity to Real World	Exp. Similar Real World	Similarity Imag. World	Excessive Realims	Photorealism	Similarity to Photo/Video
LG & iPhone	0.2456	0.2838	0.3638	0.5000	**0.0369**	0.3388
LG	0.2540	0.4196	**0.0400**	0.3233	**0.0397**	0.5000
iPhone	0.3706	0.1546	0.2392	0.3383	0.2594	0.2687

Table 15 – Visual Realism: Island vs Cave display (p values).

Analysis: The overall outcome confirmed the high level of visual realism experienced by the users on both environments. The significant higher values related of the contribution to photorealism in the cave is surprising because of the nearly monochromatic appearance, which was supposed to undermine it, and of the portrayed closer distance to objects, expected to make it easier the discovery of deformations [148]. The higher LG's pixel-density appear to counterbalance the above aspects by providing sharper images, which as earlier mentioned were well appreciated on closer objects, and the cave has plenty of them.

5.6.2.2 Presence

Results: Both environments scored high the *sense of being there* (P1), *surrounding reality* (SP1), *feeling presence in virtual space* factor (SP3) and *feeling present in virtual environment* (SP5). The SP1, SP3 and SP5 saw significant higher values in the island when merging scores of both displays (*p*=0.0273, *p*=0.0273 and *p*=0.0368 respectively), and on LG display only (*p*=0.0436, *p*=0.0382 and *p*=0.0141 respectively). Low were the scores of the *perceiving pictures only* (SP2), and medium were the scores of the *sense of acting there* (SP4), with the cave scoring SP4 significantly higher than the island on all display combinations (*p*=0.0096 on both display data, *p*=0.0429 on LG display only, and *p*=0.0487 on iPhone display only). Table 16 shows all *p*-values.

	Being There	Surr. Reality	Pictures Only	Pres. Vir. Space	Acting There	Pres. in VE
LG & iPhone	0.4034	**0.0273**	0.3763	**0.0273**	**0.0096**	**0.0368**
LG	0.4362	**0.0436**	0.3117	**0.0382**	**0.0429**	**0.0141**
iPhone	0.2840	0.1618	0.5000	0.1567	**0.0478**	0.3287

Table 16 – Overall and Spatial Presence: Island vs Cave display (p values).

The *real world awareness* and *unawareness* (INV1 and INV2) scored high (median 3) and it scored similarly in both environments. The *attention to surrounding reality* (INV3) scored low with a tendency to significant higher values in the island on both displays combined (*p*=0.0569). The *attention to the VR world* (INV4), scored high on both environments with the island performing significantly higher on both displays (*p*=0.0001), on LG display only (*p*=0.0016) and on iPhone display only (*p*=0.0074). Table 17 shows all *p*-values.

	Real World Aware	Real World Unaware	Att. Reality	Att. Vir. World
LG & iPhone	0.3461	0.2751	**0.0569**	**0.0001**
LG	0.2960	0.3932	0.1289	**0.0016**
iPhone	0.5000	0.2845	0.1337	**0.0074**

Table 17 – Involvement: Island vs Cave display (p values).

Analysis: Both environments gave users strong *overall presence* and *spatial presence*. The significant better performance in the island compared to cave on SP1, SP3 and SP5, was clearly triggered by the LG display. If we add the considerations made in the previous section on SP1, SP3 and SP5 factors (LG significantly higher performance in the island only), we observe once again it is the combination island – LG display that triggers the difference.

With the support of users' comments, we can conclude that the island's enhanced sense of presence is due partially to the wider environment-views (confirming [107]), and partially to the good lit and wide-color spectrum (confirming [149]); whereas the cave's enhanced *sense of acting* seems mainly geared by the close-distance views calling for head rotation to discover the environment.

Interestingly, both in case of the island and cave, it is the use of the LG display that makes the differences significant. Looking at the display specs we could argue it is the higher *pixel-density* triggering the difference.

The effective sense of isolation from the surrounding space the HMD provides is again confirmed on INV1 and INV2, whereas the tendency to significant higher performance on INV3 and INV4 in the island, appears due to the island's richer scenario.

Looking at the overall outcome for environment and display, it can be asserted it is the content playing a major role compared to display. E.g. the island's colorful landscape results more appealing to viewers than the cave monochromatic views, which affects *involvement* and most of the *spatial presence* factors.

5.6.2.3 Emotions

Results: The scores in both environments were high for *happiness*, *enjoyment* and *surprise*. In the island scores were high on *relaxation* only, whereas in the cave scores were high on *scariness* and *anxiety*. The scores were low in both environments on *anger*. The most significant difference is on *relaxation* with island performing higher than cave (p<0.0001 on both and on each display). The island also performed significantly higher on *enjoyment* on both displays and LG only (p=0.0488 and p=0.0279 respectively), whereas the cave performed significantly higher on *anxiety* on both displays (p=0.0031) and on iPhone only (p=0.0071). Table 18 shows all p-values.

	Happiness	Enjoyment	Relaxation	Scariness	Sadness	Anxiety	Anger	Surprise	Back Reality
LG & iPhone	0.1773	**0.0488**	**0.0000**	0.0651	0.1078	**0.0031**	0.3402	0.1044	0.3491
LG	0.3561	**0.0279**	**0.0000**	0.1576	0.2474	0.0759	0.2906	0.2068	0.3327
iPhone	0.1787	0.3383	**0.0000**	0.1303	0.1410	**0.0071**	0.5000	0.1638	0.4571

Table 18 – Emotions: Island vs Cave display (p values).

Analysis: Emotions were largely triggered in both environments. The main reasons why the island scores significantly higher on enjoyment and relaxation were identified by users as due to its warm colors. This goes along with [110], addressing contribution of lighting in realistically portrayed virtual environments. As for sadness, scariness and anxiety, we noted higher values in the cave. Fifteen of our users (75%) commented these three types of emotions were elicited by the nearly monochromatic scenes and lack of daylight. It was also noted that scariness and anxiety appeared further enhanced by the cave's closer distances (e.g. walls and stones) triggering more rapid head-movements than in the island.

We observe a similar general trend between the overall presence (P1) and the emotion's happiness and enjoyment factors, which is supported by many literature works [150], [83], [151], [146], [108].

5.7 Results and Analysis: Sense of Place

This section analyzes results focusing on the role played by place familiarity. Scores were gathered from both unfamiliar and site-familiar users looking at two different environments (island and cave) through the same display (iPhone). Referring to what is described in section 5.4.3 this means: C1 vs D1, C2 vs D2 and C1+C2 vs D1+D2. Figure 39 and Tables 19 through 21 show results.

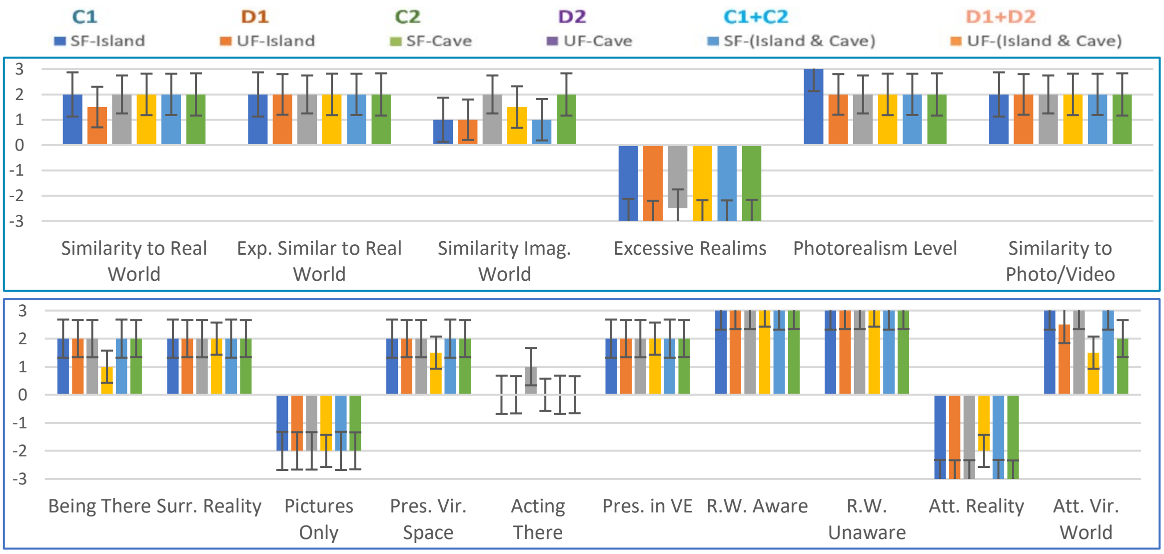


Figure 39 - Outcome for User Study 4: Visual Realism (top-row) and Presence (bottom-row). Indicated values are median and standard error. Site familiarity is indicated by SF, whereas unfamiliarity to the site is indicated by UF. The two environments are indicated as Island and Cave.

5.7.1 Human Factors

5.7.1.1 Visual Realism

Results: Users scored realness high-level (i.e., high scores on all factors except for excessive realism). There were no significant differences except on experience similar to real world, where site-familiar users scored significantly higher in the island (p=0.0458). Photorealism scored high with site-familiar users scoring it significantly higher than unfamiliar in the island (p=0.0285). Table 19 shows all p-values.

	Similarity to Real World	Exp. Similar Real World	Similarity Imag. World	Excessive Realims	Photorealism	Similarity to Photo/Video
Isl. &Cave	0.2538	0.1588	0.1628	0.2885	0.2202	0.2024
Island	0.2975	**0.0458**	0.2293	0.1792	**0.0285**	0.2577
Cave	0.2101	0.2718	0.0964	0.3978	0.4120	0.1470

Table 19 – Visual Realism: Site-Familiar vs Unfamiliar (p values).

Analysis: Visual realism scored high-level overall. It is of great interest and surprising to note the high-scores site-familiar users provided when observing the two environments. This is noted because the judgement of site-familiar users is based on correct environment knowledge, e.g. trees' and rocks' shapes. Therefore, they should be the best in spotting to unnatural appearance due to mismatches and deformations and therefore lowering their scores. It is also interesting to note that site-familiar and unfamiliar users' scores are generally similar.

Concerning photorealism, we know from the literature that it contributes to distance perception and depth-impression both for those familiar and unfamiliar to a place [116], [152], [153]. We can then observe that high photorealism seems to further support those site-familiars. However, this only happens in the island case. Users referred the island's higher scores were due to its image characteristics (illumination and colors) providing a more convincing effect. This would also justify the significant higher scores the same users give on the same environment to the *experience similar to real world* factor.

Overall, the comparison between site-familiar and unfamiliar users gave an outcome opposite to supposition. We expected unfamiliar users would overlook deformations or wrong details in scene elements, as they would not know the actual look of things, e.g. they would not notice trees that look taller or rocks with deformed shapes, whereas site-familiar users would be more critical. Rather, site-familiar users generally gave higher scores than unfamiliar users, and their comments were more appreciative. They typically commented seeing very realistically looking environments.

5.7.1.2 Presence

Results: Site-familiar users scored presence generally higher than unfamiliar users. The scores of site-familiar users on *overall presence* (P1) were significantly higher than unfamiliar users in the island (p=0.0333), and with a tendency to significant higher difference when considering both environments' scores combined (p=0.0599). We found no significant differences on *spatial presence* (SP1-SP5).

Site-familiar users scored *attention to surrounding reality* (INV3) and *attention to VR world* (INV4) significantly higher in the cave (respectively p=0.0235 and p=0.001). Tables 20 and 21 show all p-values.

	Being There	Surr. Reality	Pictures Only	Pres. Vir. Space	Acting There	Pres. in VE
Isl. & Cave	**0.0599**	0.2286	0.4575	0.1928	0.2469	0.1591
Island	**0.0333**	0.2719	0.5000	0.2189	0.1322	0.1690
Cave	0.0864	0.1853	0.4150	0.1668	0.3616	0.1491

Table 20 – Overall, Spatial Presence: Site-Familiar vs Unfamiliar (p values).

	Real World Aware	Real World Unaware	Att. Reality	Att. Vir. World
Island&Cave	0.2987	0.2729	0.2166	0.1440
Island	0.2960	0.3978	0.4096	0.2869
Cave	0.3015	0.1480	**0.0235**	**0.0010**

Table 21 – Involvement: Site-Familiar vs Unfamiliar (p values).

Analysis: The generally higher scores of site-familiar users seemed connected to the more enthusiastic attitude these users had. According to their comments, the positive attitude came by seeing realistic visual reproductions of places they knew well, which brought memories. This is in line with some literature work indicating a correlation between familiarity and *presence* [154], [155].

Interestingly, the significantly better performance of site-familiar users when observing the island did not occur in the cave for P1, because unfamiliar users scored it higher. Based on users' comments, those higher scores were given because of the stronger depth impression the cave delivered, which enhanced *presence*.

5.8 Conclusions

This research phase investigated the effect of photographic realism in VR when observing real places through a VR headset. We focused on the role played by: pixel-density (screen); illumination and object distances (location); previous knowledge (sense of place). Experimental data were gathered by interrogating users on the effect of a number of display and human factors (image lighting, visual realism, presence and emotions), which were presented and analyzed according to the research questions. The main outcomes are summarized below.

- *Screen*. The users appreciated the quality of both screen displays. They felt that image lighting factors, such as color, lightness, vividness and sharpness, contributed to realistic viewing. The contribution of higher pixel-density was positively felt and it prevailed over better lighting specs, leading to a significant improvement in some realness and spatial presence factors.
- *Location*. The locations' environment showed to affect the felt contribution to realistic viewing provided by pixel-density, lightness, color, vividness, and photorealism. It relevantly influenced spatial presence and in particularly the sense of acting there (in the visualized world), as well as involvement in terms of attention to VR world. Some emotions were clearly elicited. They were enjoyment and relaxation (while observing the island) and anxiety (while observing the cave). Both location's illumination and distance to objects appeared to contribute towards depth impression, respectively in the island and cave. The cave scored a significantly higher depth impression because of the binocular depth-cues induced by the close distance to objects.
- *Screen-Location Combination*. The location generally played a stronger role than the display, which proved that good quality displays are "transparent" to scene content. Nonetheless, the display characteristics were still able to further enhance image-lighting, spatial presence and some elements of visual realism and emotions, leading to significant score differences between one specific screen-location combination and all the others. This was the case for the LG-Island and the LG-Cave combinations. In case of emotions, the higher pixel-density display specifically enhanced enjoyment in the island and anxiety in the cave.

- *Sense of Place*. The effectiveness of our systems was confirmed by users knowing well the place scoring high visual realism and presence. Interestingly, place-familiar users scored some factors even significantly higher than users that did not know the place, which proved that previous knowledge can positively enhance perceived realism and presence, e.g. by bringing memories. This represents a fascinating aspect worth future investigation.

The performed studies proved the overall effectiveness of the visual experience provided by 3-D omnidirectional photorealistic images, observed through a VR-headset, and visualized according to the proposed system concept. Users' scores were generally high for *visual realism*, *presence and emotions*.

We deem the outcome of both experiments was positive, particularly if we consider the limitations in terms of static images and a choice for system elements constrained by the use of off-the-shelf devices, including camera and visualization systems.

6. The True-Dimension Advantage: Depth Perception, FOV and View Adaptation

This chapter addresses depth perception and distance estimation. It starts with presenting a user study that collected user's depth impressions and tested user's distance-estimation accuracy. It then describes some developed techniques to reduce deformations and improve depth perception in photo-based VR. The last section presents the design and setup of a new procedure to reduce and counterbalance the arising viewing distortions.

6.1 User Study 5: Photo-Based Depth Perception and Estimation

6.1.1 State of Art

Vision is the main human sensory modality, the sensorial input we believe the most, and with a demonstrated ability towards perceiving object location in 3D-VR space [112], [44]. [156], [157]. Visual realism has therefore also been studied in terms of its contribution to distance perception and estimation, which are claimed to be key elements in realistic VR viewing [23], [127]. Image quality and its relation to depth perception had gathered wide interest in the last decade, thanks also to the development and marketing of 3DTVs.

There has been a research focus on picture and depth quality and their effects to naturalness and presence. The work of Kuijsters et al. [87] discusses contribution of S3D to naturalness and depth perception. The work of Li et al. [158] focuses on the role of the FOV towards distance judgments, whereas the work of Li et al. [159] focuses on the role of human peripheral vision. Literature works have also focused on depth perception and distance estimation, and established connection between stereoscopic-3D viewing and user's performance in a number of applications of VR for telepresence and teleoperation [106], [160], [161].

6.1.2 Research Questions and Measurements

6.1.2.1 Research Questions

- *Depth Perception*. How depth impression and contribution vary for different types of environments?
- *Distance Estimation*. How distance estimation varies for different types of environments?

The literature has shown that the portrayed environment characteristics and display factors can affect distance estimation.

6.1.2.2 Evaluation Measurements

The aim is to answer the above research questions (representing independent variables) by asking users to judge on depth perception and distance estimation (dependent variable).

The use of subjective ratings and measurements is proposed. The ratings are related to depth impression, whereas measurements are related to the perceived distance to objects.

Depth Perception

As seen in previous works, we ask users generic questions about the delivered sense of three-dimensional appearance. We also ask for users' opinion on contribution of depth impression towards other factors.

- Depth Impression. (a) Overall depth impression (delivered sense of three-dimensional appearance); (b) Speed to get 3D impression.
- Depth Contribution. (c) Depth impression contribution to realism; (d) Depth impression contribution to emotions; (e) LSC to 3D (lights, shadows and colors contributions to depth impression).

Distance Estimation

Distance estimation and perceived depth have been measured in different works through quantified judgements reported by users [89], [114]. The methods typically rely on either interactive procedures or motionless observations.

In case of interactive procedures, distance is estimated by asking users to perform specific actions, e.g. when investigating FOV and minification in VR images. Some works estimate distances through blind walks [105], [23]; others through directed walks [127]; and through triangulated walks techniques [113].

In case of motionless observations, distance is estimated by asking users to perform observations statically. Relevant examples based on image comparison include the works of Hibbard et al. [129], which associates eye-disparity increase and realism of S3D views to depth judgement, and Baek et al. [162], which investigates the role played by different displays and related parameters (FOV, resolution, brightness, S3D, camera distance) to distance estimation.

In our assessment, we are most related to the case of motionless observations, being our observed scenery static with users not allowed to move except for turning their head around. We therefore follow what is a common approach for distance perception in this type of observation, which is based on motionless depth judgements. We ask users to estimate egocentric distance to 6 pre-selected scene elements and relative distance between 5 of them. We then estimate the average relative error as in [105] and depth perception accuracy based on Baek's formulation [162].

For our distance estimation, we consider a range where humans (under direct viewing) can generally benefit from binocular vision (0.3-10 meters) [163]. This is relevant because of its demonstrated positive contribution to realism and presence. The proposed assessment compares users' estimates to ground truth and between the two environments.

Distance Estimation: (a) Egocentric distance estimation to 6 scene-elements; (b) Relative distance estimation between 5 scene-elements.

6.1.3 Experiment Design

A within-subject study was conducted with 20 test-users who did not have any previous knowledge of the environment. Participants' age ranged between 22 and 53, with an average of 28. The study had the conditions described below. The two chosen environments were the same described in section 5.4.5 (island and cave). The chosen display was the *LG G-series*. The test organization and procedure are described in figure 40. It followed the experimentation steps described in section 5.4.4 with the inclusion of one additional step called *Estimation*.

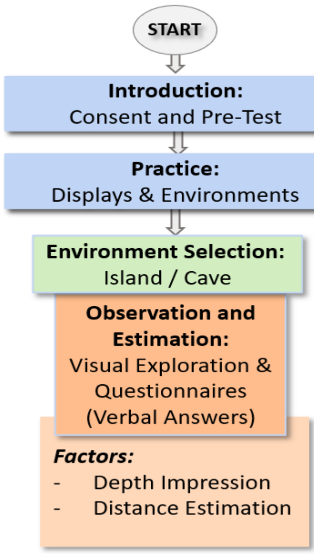


Figure 40 –Main experimentation steps

In the *Estimation* step participants were asked to turn their head to see a specific portion of the panorama (at a pre-designed position and orientation). They could not walk, but they could turn their head around. Participants were then requested to verbally provide a set number of measurements (in meters), each including an integer and one-digit fractional part. They could observe for how long they needed.

The distances asked to be estimated were between 2 and 9.2 meters. The locations of the 6 observation points were on each environment at approximately the same distance and orientation from viewers. The set locations also followed the same distribution pattern. Positions for distance estimates were randomly queried. The range of distances were set similar to that used in the work of Interrante et al. [105] to allow for data accuracy comparison. Figure 41 shows the set views on each environment and table 24 shows the ground truth distances.

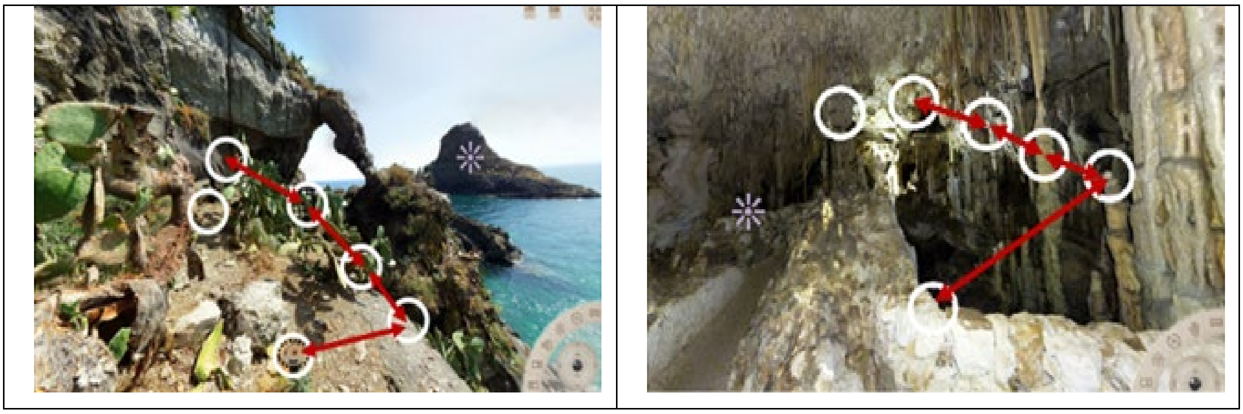


Figure 41 – This island' and cave' set scenarios for users' distance estimation: egocentric distance to the center of while circles; relative distance among the circles (red arrows). The white circles and red arrows are only for reader's comprehension and were not shown to users. Users held a small red cross indicating the locations.

6.1.4 Results and Analysis

6.1.4.1 Depth Impression

Results: Figure 42 shows the outcome of our Depth Impression questionnaire. The test was only conducted on the LG display for practical issues. The overall depth impression scored high on both environments, with the cave performing significantly (p=0.0377).

Medium to high values were the scores of the other factors with no significant differences. We noted the depth contribution to emotions was generally higher in the cave. The LSC to 3D (benefits from light, shadows and colors towards depth impression) scored high in both environments, with the island performing significantly higher (p=0.0171). Table 22 shows all p-values.

	Overall DI	Speed to DI	DI Realism	DI Emotions	LSC to 3D
LG	**0.0377**	0.1394	0.1423	0.0967	**0.0171**

Table 22 - Depth Impression: Island vs Cave display (p values).

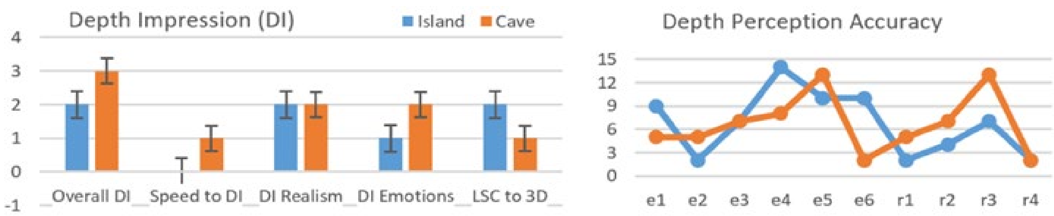


Figure 42 – Left: Outcome for: Depth Impression; indicating median and standard error values. Right: Depth Perception Accuracy in % is estimated as in [75]. A 5% error is considered neglectable.

Analysis: The overall depth impression factor scored high in both the environments, despite their difference in terms of light-intensities and distance-range. This shows our system successfully provides three-dimensional impression. Users commented to get a greater depth impression on the closer portrayed objects (e.g. cave stones and walls), which is in line to what happens with real environment observations [164], [165]. It confirms the relevant contribution of binocular depth-cue on short-medium distances (0.3-10 meters (Read, 2015)), which occurs more relevantly in the cave because of the more objects at closer distances. Users indicated the strong depth impression contributed to enjoyment in both environments (and to perhaps to scariness too as commented regarding emotions). This is a fascinating aspect worth further studies.

Concerning the LSC to 3D factor, the illumination clearly played a role towards 3D impression, which is in general agreement with literature works on the contribution of light, shadows and colors to depth perception [166], [149], [99]. The LSC to 3D significantly higher performance in the island was according to the 90% of users, due to the many colors. We deem that colors particularly supported the monocular depth-cues induced at the higher distance-range (over 10 m.), compensating for the lack of binocular depth-cues [163]. Users also indicated that shadows were felt in the cave as main contributors to the perceived depth impression.

6.1.4.2 *Distance Estimation*

Results: Figure 43 shows users' performance for egocentric distance estimation. The test was only conducted on the LG display for practical reasons. Both in the island and cave, we could observe higher accuracy in our intermediate locations (e3, e4, e5, r2, r3). Table 24 shows ground truth values for island and cave locations.

The average relative error was below 5% for island's r2 and r3 and cave's e4 and e5, and only 2% for the island's e4. The errors for island's r3 and e4 were significantly lower compared to the analogous cave's locations (p=0.021, p=0.045), whereas the error for cave's e5 (compared to island's e5) was with a tendency of being significantly lower.

The average relative error was between 5% and 8% for island's e1, e3, e5, e6 and cave's e3 and r3; whereas it was between 12% and 19% for island's e2 and cave's e1 and e6. The errors for island's e1 and e6 were significantly lower, and that for island's e2 significantly higher, compared to cave's equivalent locations. The p-values were respectively p=0.019, p=0.002, p=0.020). There were otherwise no significant errors. Table 23 shows all p-values.

	e1	e2	e3	e4	e5	e6	r1	r2	r3	r4
LG	**0.019**	**0.020**	0.293	**0.045**	0.058	**0.002**	0.124	0.060	**0.021**	0.075

Table 23 - Depth Estimation: Island vs Cave display (p values).

e1		e2		e3		e4		e5		e6		r1		r2		r3		r4	
2.0	2.0	3.5	3.1	4.3	4.0	5.0	5.4	6.1	6.2	9.2	9.2	2.4	2.1	1.2	1.0	2.1	2.0	4.8	4.6

Table 24 - Locations' ground truth value for egocentric distance and relative distance estimations (in meters).

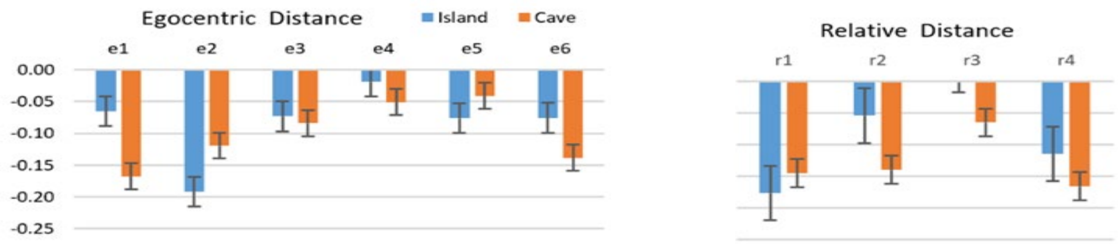


Figure 43 - Average relative error (in meters) for egocentric distance and relative distance estimated by users for positions e1 through e6, and for positions between r1-r2, r2-r3, r3-r4 and r5-r6.

Analysis: The average relative error is contained, and it is also comparable to that measured in the work of Interrante et al. [105]. The above facts imply the proposed acquisition and visualization settings well support realistic distance estimation. The estimates' accuracy is overall comparable in both environments despite the differences in illumination and objects' distance. This indicates acquisition and visualization settings play a greater role than environment characteristics. Figure 7 diagrams also show users typically underestimated distances (negative error values), which is typical when observing synthetic scenes through an HMD [23], [158].

The good accuracy when observing the island (with the higher number of significantly better values), is sustained by many users' comments. They indicate that the cave's represented stalactites and stalagmites with relatively smooth surfaces of unknown size and shape, make more difficult to comprehend their precise 3D locations. On the other hand, the outdoor island's views portraying more complex and articulated objects' shapes, such as the green multi-directional prickly pears' blades, provide a well contrasted object's appearance that makes easier comprehend object's 3D positions and orientation. Furthermore, the island observed points represent smaller surfaces than in the cave, which also supports distance-estimation accuracy [162].

6.2 Focused Developments: FOV and Concentric Spheres

Further to assessing depth perception and distance estimation, it was decided to examine the main processing steps (image acquisition, image processing and image rendering) and then focus on the analysis of some relevant elements. We believed this would help devise new ways to improve realistic depth perception in photo-based VR.

A number of focused developments and assessments were proposed to address visual realism and depth perception. They were:

1. 3D-360° Photo and Video Acquisition and Visualization for HMD;
2. VR Headsets Display Field of View Estimation;
3. Mix Reality on smartphone-based HMDs and player with variable FoV;
4. HMD Zooming Player Implementation and Concentric Spheres.

6.2.1 3D-360° Photo / Video Acquisition and Visualization for HMD

Aim: The aim of this first focused development was to get acquainted with some hardware and related software processing. In particular: photo and video cameras, VR headset display and image acquisition, processing and visualization software.

Research Questions: How is it possible to capture 3D 360-degree photorealistic environments and view them through VR viewers in the most realistic possible way? How to convey the acquired view to the HMD display, to make users feel as if they are present in the observed environment?

Approach: The VR equipment available at the UH VR and Robotics Lab was used to learn and experience VR theories, VR technologies, and the relevant visual human physiology, The approach was "learn by doing". It involved learning and experiencing and VR programming as well. Figure 44 illustrates the available VR equipment.

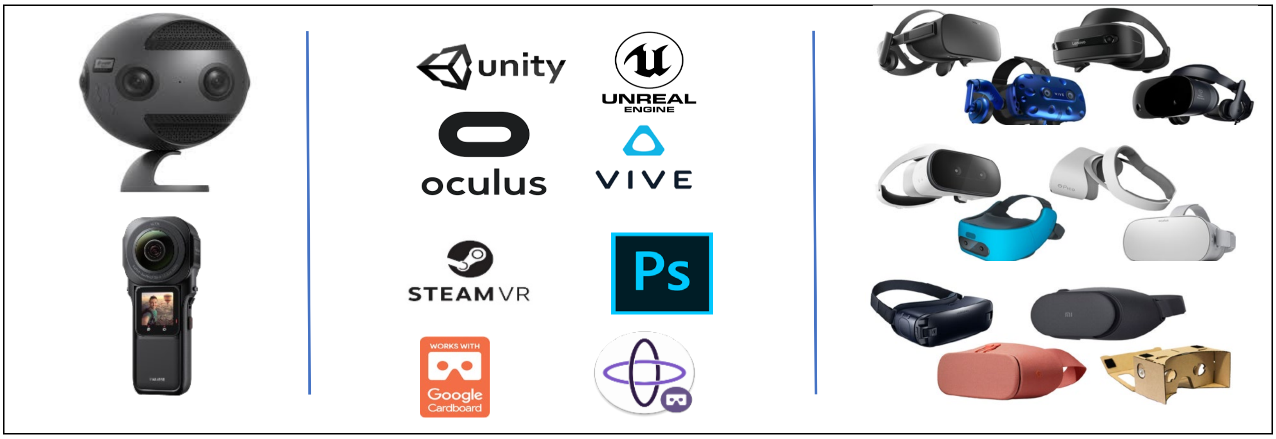


Figure 44 – VR equipment.

Testing Environment: All considered testing environments were represented by high-definition photographs. The chosen scenes were of different type (indoor and outdoor), size (small and large spaces) and illumination (natural and artificial). Figure 45 shows examples of the testing environments.

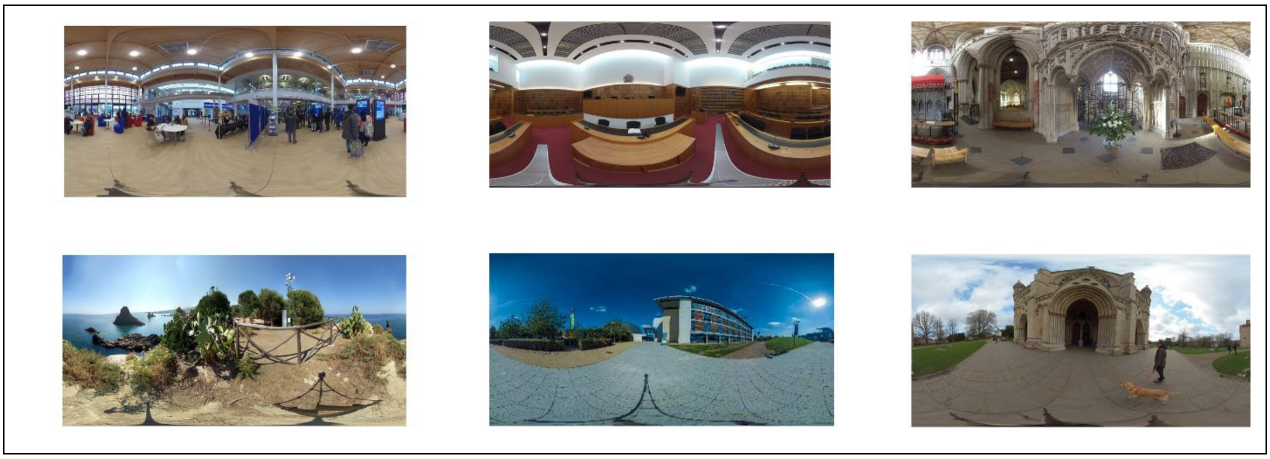


Figure 45 – Photorealistic testing environments.

6.2.1.1 Image Acquisition

The Camera *Insta360 Pro* allowed for acquisition of three-dimensional images with a 360 degrees' viewing angle and a resolution up to 8K. The acquired images could be observed three-dimensionally through the VR headset's display. Thirty different environments were acquired both representing indoor and outdoor sites. Figure 45 showed six of them.

A few cameras were used for image acquisition, included the *Insta360 X3*. Nonetheless, most of the photos were captured using the latest generation *Insta360 Pro* camera. This camera system had 6 individual cameras that capture the entire environment and provide a merged view represented as two stereoscopic 360° images (one for the right eye and one for the left side). Those 360 images can directly be seen through the VR headset. The entire 360 image can be overviewed by turning the head.

As follow up to hardware acquisition and visualization there was a phase of practising and learning the Unity VR software. This allows for mapping the acquired images into a 3D spere, which represents the 3D model that will be used for view rendering.

6.2.1.2 Image Visualization through VR Headsets

The acquired 360 image is then mapped in to a sphere following the procedure introduced in section 2.1.2.

The process of visualizing the acquired image was developed and tested on the 6 different VR headsets listed below. They were of different type, including PC-powered, standalone and smartphone-based VR headsets. The evaluated VR headsets are listed below and depicted in Figure 46.

- Oculus Quest 2 (standalone);
- Oculus Quest (standalone);
- Oculus Rift (pc-powered);
- HTC VIVE Pro Eye (pc-powered);
- Samsung Gear VR with Samsung S8 phone (standalone - smartphone-based);
- Card-Board VR with Samsung S8 phone (standalone - smartphone-based).

Figure 46 – VR Headsets.

Developing for the above-listed headsets required different techniques for code development and deployment. There are many ways to visualize 3D-360° images on a headset. The investigated and developed solutions are listed below and described in the following paragraphs.

- VR Browser
- Embedded VR Player
- Independent VR Player
- Unity player with Skybox Shader
- Unity player with Concentric-Circles procedure.

VR Browser

The acquired image was processed by a software for 360 images, we used the *Kolor Panotour Pro* software, which allowed users to visualize 360 images through the VR headset's browser. The images were uploaded on a web server. This technique therefore required the presence of an Internet connection. This solution worked well on the standalone headset *Oculus GO,* and on the smartphone-based headsets *Samsung Gear VR* and *CardBoard VR*.

The acquired images were related to different university websites. The images became part of a database, which was visualized and managed through the browser. Immersive VR views could also be rendered through the developed software.

Figure 47 shows images of the VR browsers, and Figure 48 shows the developed website, which visualises and manage the image-database through the browser.

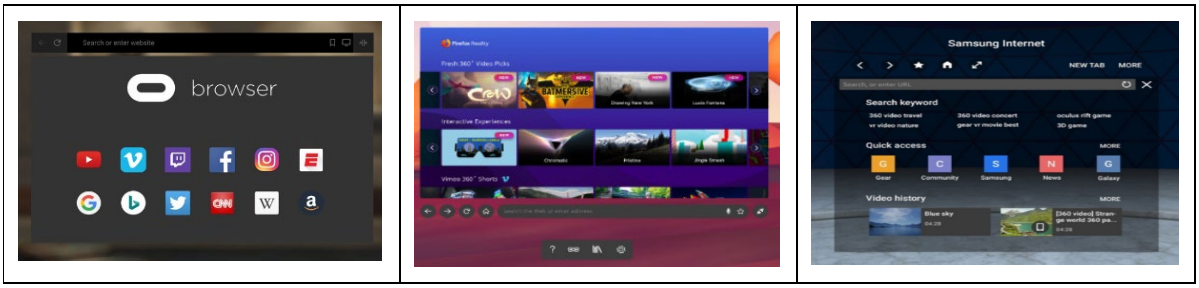


Figure 47 – Browser VR Player.

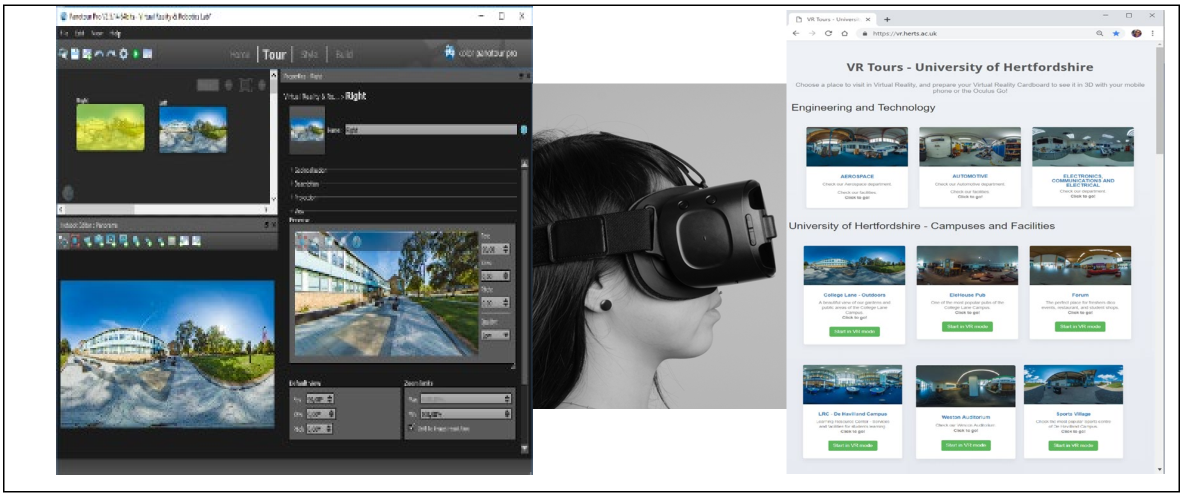


Figure 48 – The developed website, which visualises and manages the image-database through the browser. The website can be found at the following address: http://vr.herts.ac.uk

Embedded VR Player

Multimedia content could be viewed through the headset's embedded player. All headsets had their own operating system, and this had an embedded photo and video 2D/3D player. The embedded VR player made possible to watch 360° images using the headset head-tracking system. Figure 49 shows two embedded VR players related to the Oculus Quest 2 and HTC Vive Pro-Eye.

In order to view photos and videos both on a PC monitor and VR headset, a canvas tool was used, which was placed in the virtual space of the HMD. The acquired images were then either loaded on the PC Memory (in case of a PC-powered headset), or directly transmitted through a cable or by wireless connection to a standalone VR headset system.

Headsets based on smartphone with Android operative system, like the *Samsung Gear VR* or *Cardboard VR*, did not have a pre-installed 3-D Player. An Independent Player was then needed, which could be found on the *Google Store*.

Figure 49 – Embedded VR Players. (Left) Oculus Quest 2; (Right) HTC Vive Pro-Eye

Independent VR Player

The independent player functioned like for the embedded player. However, it was acquired through the web. The main difference between the two players relies on the greater flexibility the independent player possess. This allowed us to implement the zooming in/out option. This meant that the VR headset's FOV could be adjusted. Figure 50 images related to the independent player.

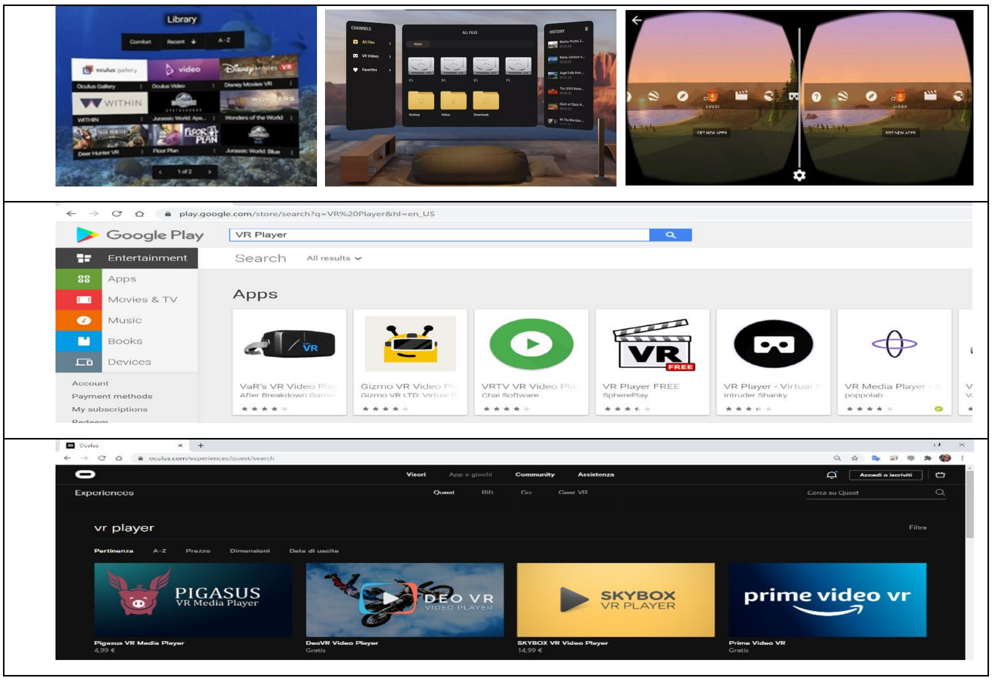


Figure 50 - Images related to the independent VR Player.

Unity VR player with Skybox Shader

The acquired images were loaded through the *Unity* software and mapped into a sphere. They were then visualized by an ad-hoc built player that uses the prefabricated *Skybox Shader* function. This player did not allow for FOV adjustment. Figure 51 shows the developed Unity VR player with Skybox Shader.

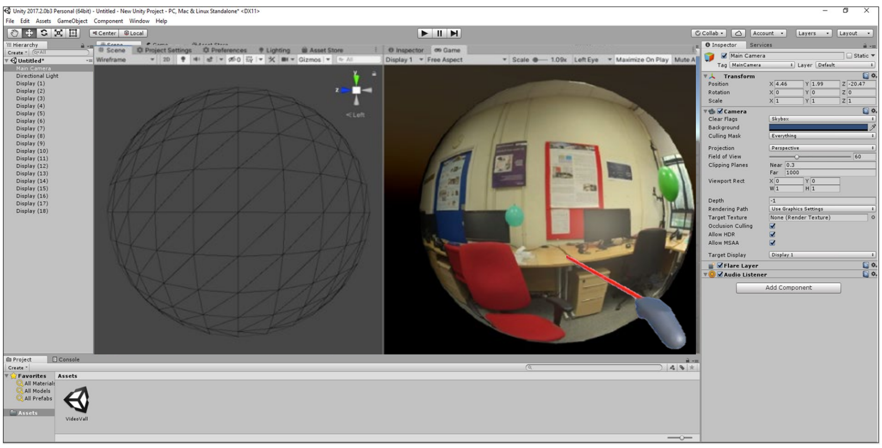


Figure 51 – The developed Unity VR player with Skybox Shader.

Unity VR Player with Concentric-Circles procedure

This is like the above-described player but the player's prefabricated *Skybox Shader* function had been removed in order to set a different procedure that would allow the user to have control over the zoom. The developed procedure for zoom-control was needed to allow the virtual camera (called *camera-rig* in Unity) to move inside the sphere, which affected the observed FOV. Figure 52 illustrates the zoom-control procedure for different virtual-camera positions and the corresponding FOVs.

The zoom-control will allow the rendering to benefit from an ad-hoc navigation procedure named *Concentric-Circles* that would make use of a specifically designed *Shader* function. The *Unity Player with Concentric-Circle* procedure is described in section 6.2.4.

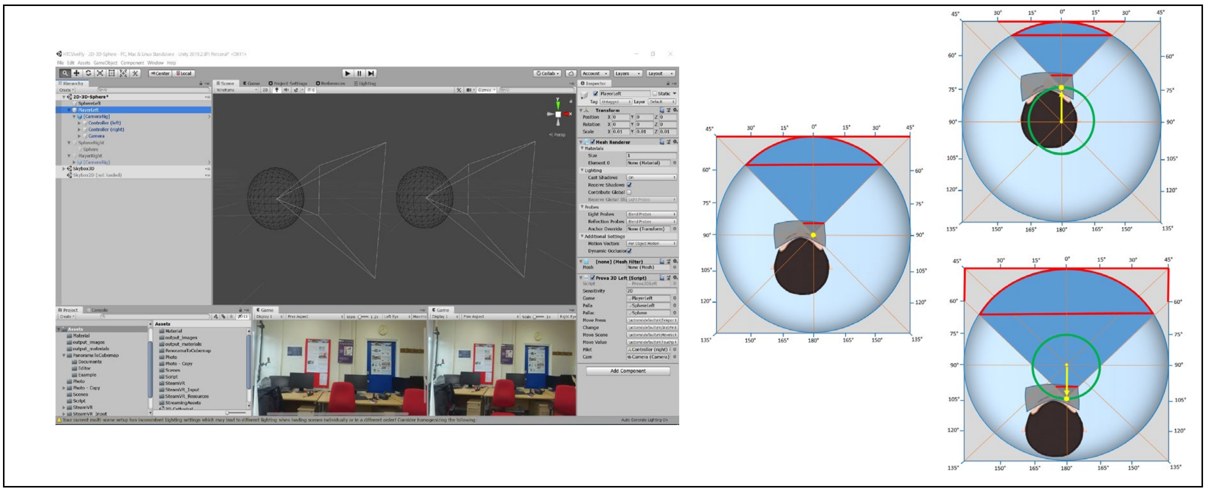


Figure 52 – Unity VR Player with Concentric-Circles procedure.

In summary, the developed players were tested on the different VR headsets indicated in the Table 25.

	Embedded Player	Independent Player	Browser	Unity player with Skybox Shader	Unity player with Concentric-Circles procedure
Oculus GO	X	X	X	X	X
Oculus Quest	X				
Oculus Rift	X				
HTC VIVE	X			X	X
Samsung Gear VR with Samsung S8 phone	X		X		
CardBoard VR with Samsung S8 phone		X	X	X	X

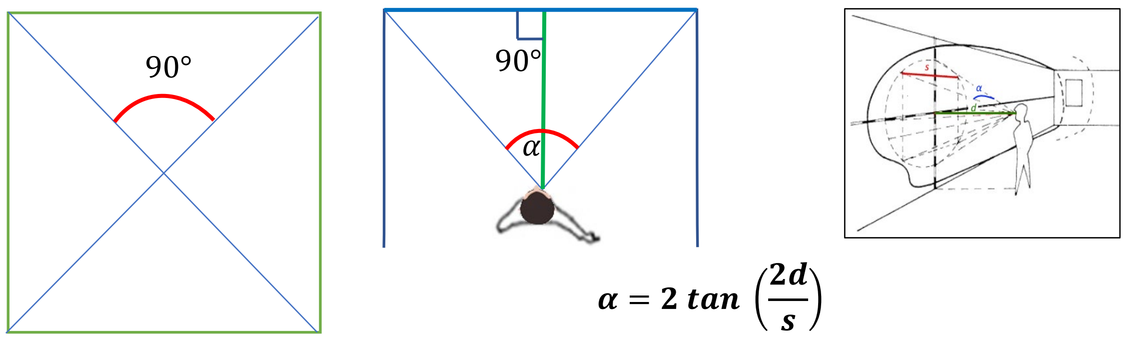
Table 25 – The tested VR Players and HMD.

6.2.2 VR Headsets FOV Estimation

Aim: The aim of the second focused development was to estimate the VR Headsets' Field of View (FOV). The reason for this assessment was that the FOV is expected to play a relevant role in providing users a sense of presence in the displayed environment, and also in terms of perceived distances.

Research Question: What is the FOV of different VR Headsets? How can we estimate it?

The FOV measurement was accomplished through display calibration. We were only interested in estimating the horizontal FOV. The VR Headset horizontal FOV could be estimated when knowing observation position, camera position during acquisition, display size and distance to viewer, and the size of the observed environment. The environment was a squared room and it was divided in four equal slices 90 degrees apart from each other to facilitate measurements and calibration. The details of the environment and related settings are shown in Figures 53 and 54.



$$\boldsymbol{\alpha = 2\, tan\left(\frac{2d}{s}\right)}$$

$$Field\ of\ View = 2 * \ tan\ \left(\frac{2\ *\ distance\ from\ display}{size\ display}\right)$$

Figure 53 – FoV Estimate.

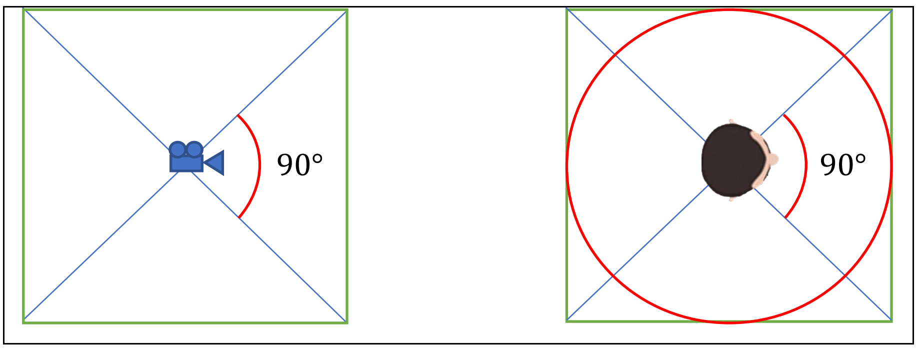


Figure 54 – Camera Acquisition. HMD Visualization.

The procedure included the action of matching the full headset's view with one side of the squared room. Figure 55 right-hand side shows the expected match. The FOV was then estimated based on the formula shown in Figure 53.

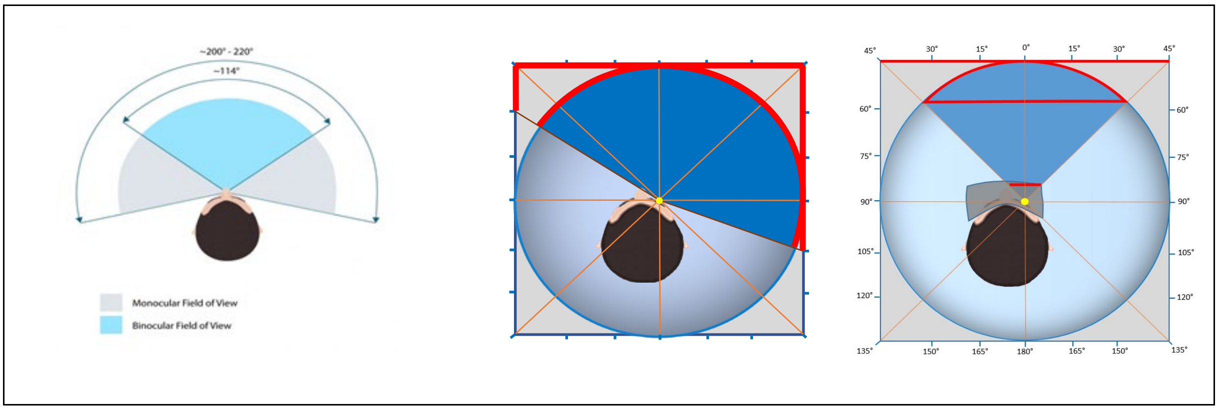


Figure 55 – Top-view describing different FOV settings. (Left) The human monocular and binocular FOV. (Center) The human right-eye FOV. (Right) The action of matching the full headset's FOV with one side of the squared room.

Testing Environment

The test-user stood at the cross section of the four slices (the room centre). The test-user's observation position was the same as the camera position at acquisition time. Figure 56 there is a schematic representation of the environment with camera and the observation positions.

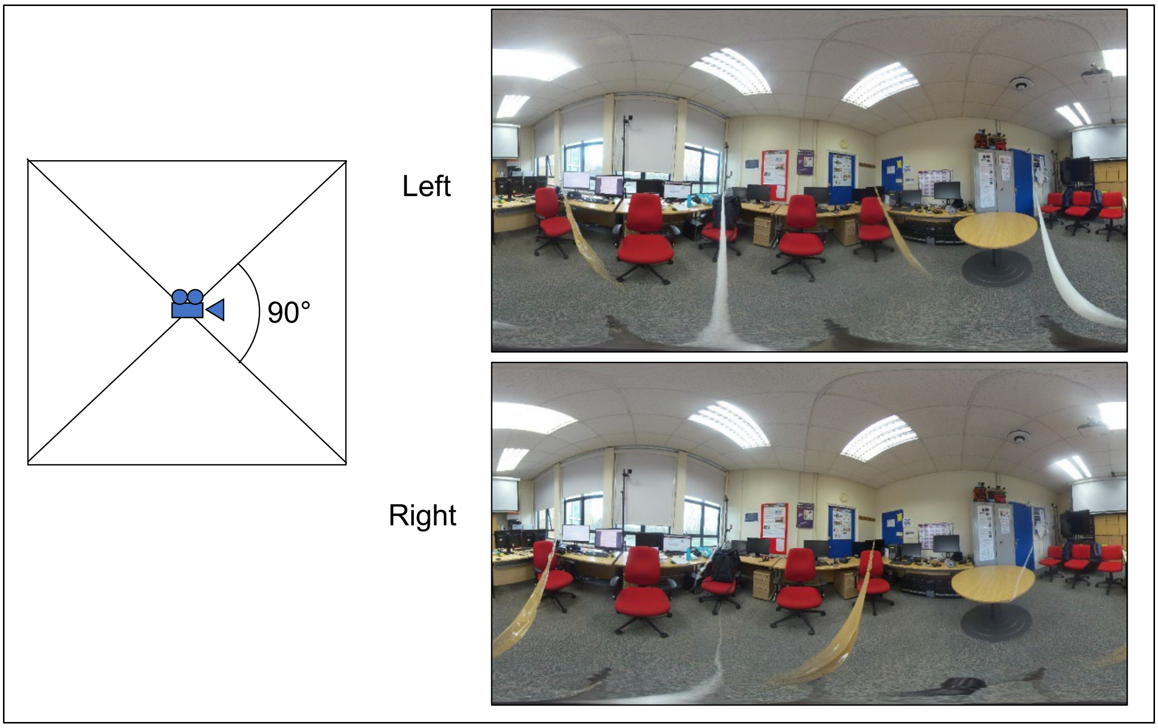


Figure 56 – (Left) A schematic representation of the environment with camera and the observation positions. (Right) The acquired 360° stereoscopic image couples.

Image Acquisition

The image to be observed through the VR headset's displays was three-dimensional and had 360 degrees viewing angle. It was taken at the intersection of the represented four slices. The room images were captured by the Insta360 Pro camera, which allowed for the acquisition of three-dimensional images with 360 degrees' viewing *angle* and a resolution up to 8K. Figure 56 right-hadnside shows the acquired images

Image Visualization

The testing Environment was observed with different VR Headsets using the methods and the players previously described. Figure 57 shows the acquired images observed through the HMD. The vertical lines represent the 4 sides of the squared room.

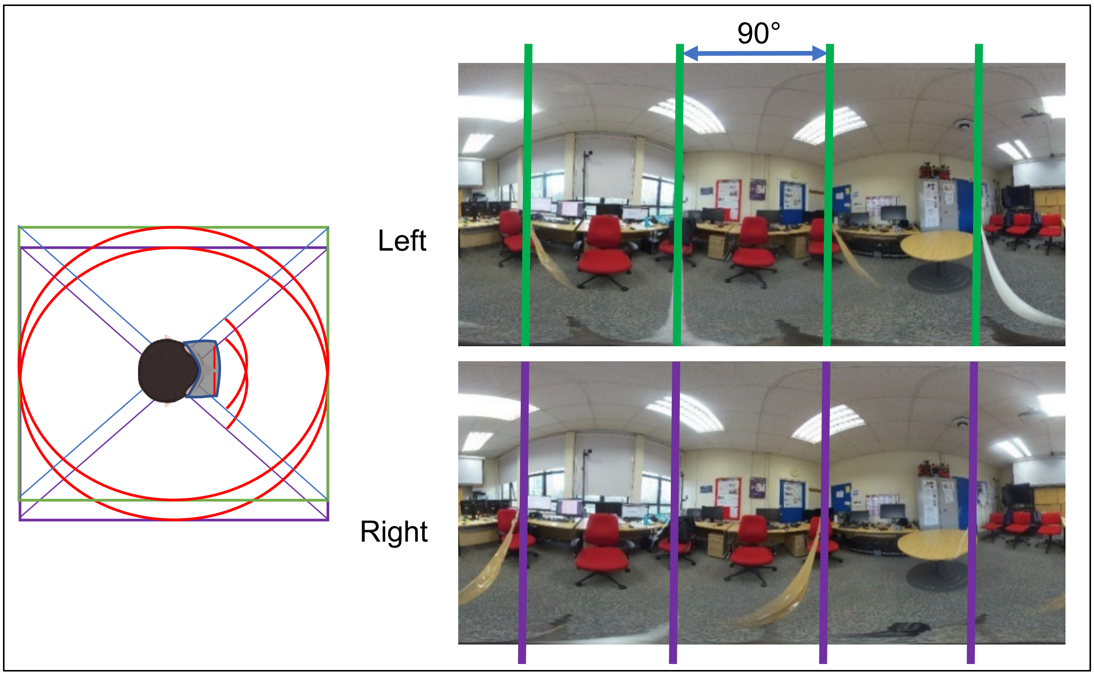


Figure 57 – The acquired images observed through the HMD. The vertical lines represent the 4 sides of the squared room.

Test Results

When observing through the headset, it often happened that the observed FOV was different from that in the headset specifications. Figure 58 shows an example when the observed FOV is 80° rather than 90°. The table 26 shows the outcome of the experiment.

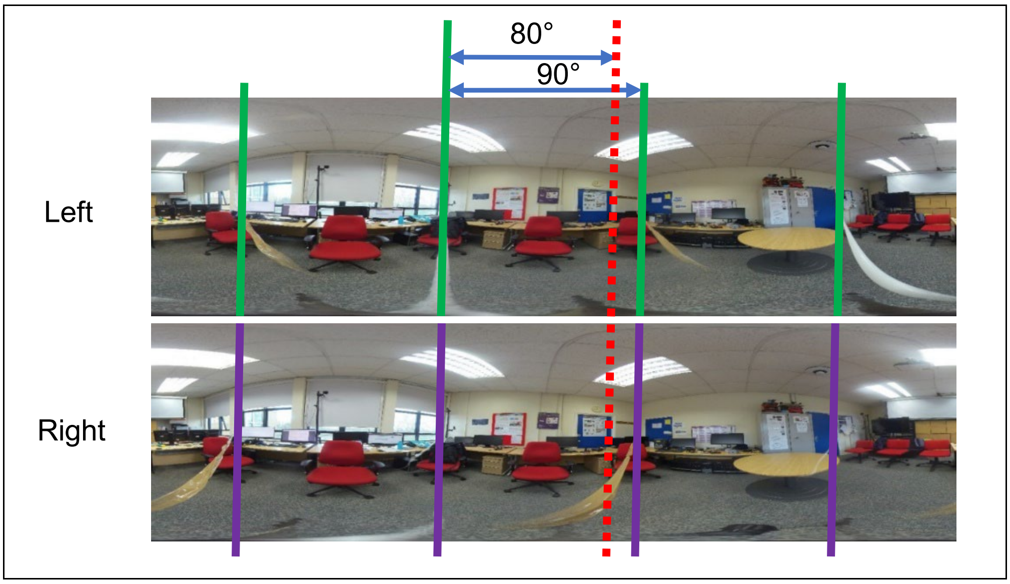


Figure 58 – The acquired images observed through the HMD. The vertical lines represent the 4 sides of the squared room. An example is shown where the observed FOV is 80° rather than 90°

FoV	Embedded Player	Independent Player	Browser	Unity player with Skybox Shader
Oculus GO	80°	80°	80°	80°
Oculus Quest	80°			
Oculus Rift	80°			
HTC VIVE	90°			90°
Samsung Gear VR with Samsung S8 phone	80°		80°	
CardBoard VR with Samsung S8 phone		30° – 200°	80°	80°

Table 26 - The headsets' measured FOV.

6.2.3 Mix Reality on smartphone HMDs and players with variable FOV

Figure 59: Mix Reality with cardboard smartphone based and player with variable FoV.

*Aim***.** The aim of the third focused development was to calibrate the headset FOV to resemble the human FOV.

Research Question: What is the correct value of a headset display FoV?

All the HMDs, assessed in the previous development, which used the various developed players, had a fixed FOV. The FOV value was generally provided by the manufacturer. Only the cardboard-type headsets, based on smartphones, had the possibility of having a variable FOV when the independent VR Player was used.

The use of the smartphone-based cardboard headset and the independent player with variable FOV were then proposed. This way we were able to measure and verify the FOV based on the environment dimensions.

For this purpose, it was necessary to develop a special player that would allow you for controlling the various display parameters of the system. This required a specially developed cardboard headset, which had been modified by removing some parts at its edges. Figure 60 shows the customized viewer, specially built to allow for simultaneous viewing of parts of the real and virtual images.

Figure 60 – (Left) The original cardboard headset viewer. (Right) The customized viewer, specially built to allow for simultaneous viewing of parts of the real and virtual images.

Having such headset allowed for comparing the image displayed on the headset display with the real environment view. The calibration was achieved when the image shown on the HMD display would precisely overlap the one seen by the necked eye. This was possible by having the real view seen through the removed headset parts which would precisely extend the image seen on the display. The obtained view was therefore a mixed reality view. It was virtual in its central part and real at its edges.

The test was carried out by observing the scene from the same position from where the scene was taken. Figure 61 shows example views with the constructed headset.

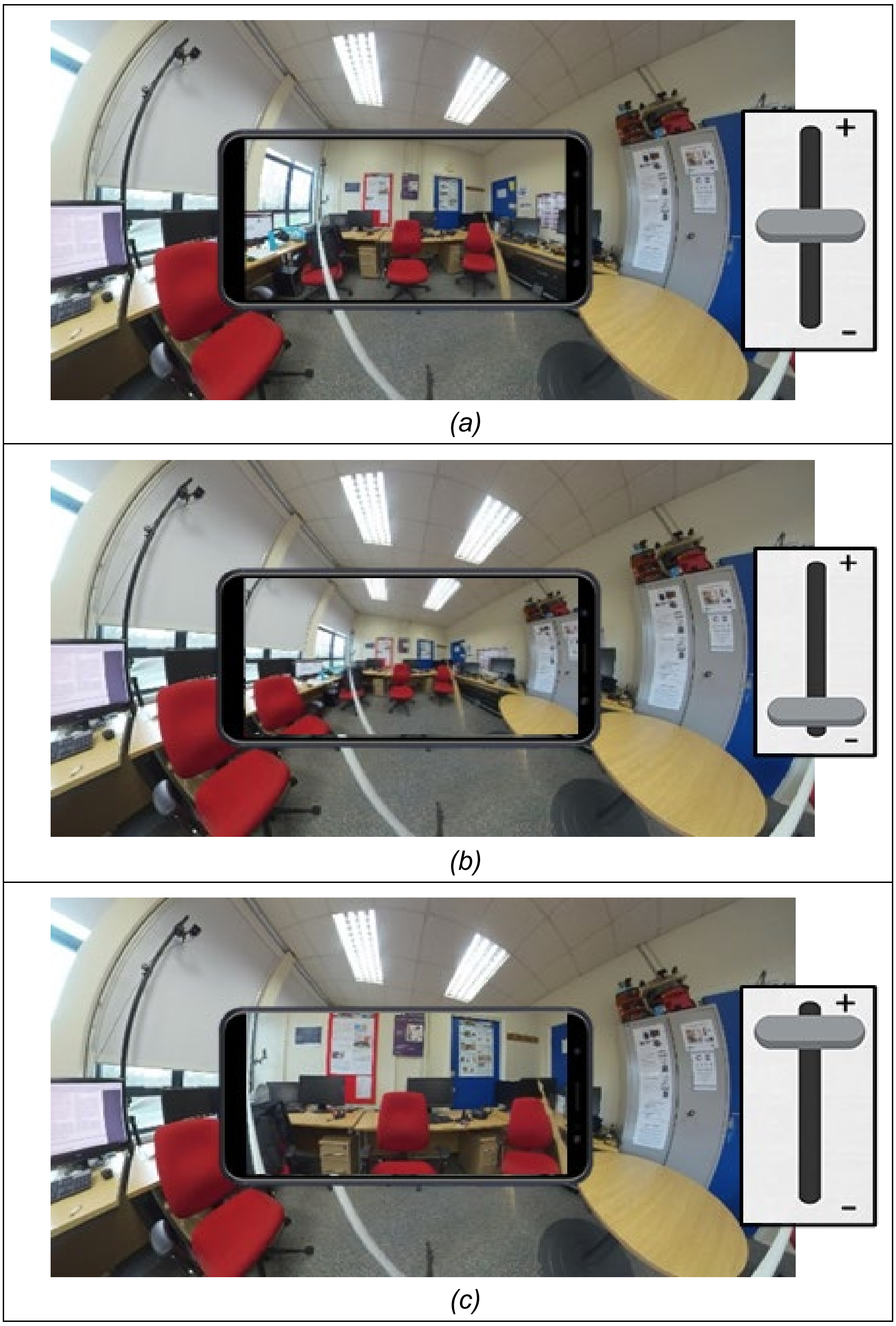


Figure 61: - Example views with the constructed headset. (a) correct FOV; (b) wide FOV; (c) zoomed FOV.

- *Test Result*: The outcome of the tests carried out for different environment settings, showed that it was possible to find the precise FoV such that it resembled the natural perceived FOV (with the necked eyes).

6.2.4 HMD Zoom Player Implementation and Concentric Sphere Navigation.

The HMD Zoom Player Implementation was made with the *Unity* software. It allowed for having a variable HMD display's FoV and the above-described calibration. This time it used the proposed algorithm named Concentric Spheres, which is below described.

Aim: The aim of this fourth development was the implementation of a player that allowed for varying the HMD FoV in order to perceive the distance of the objects in the Virtual environment similarly to the distance of the objects in the real environment, i.e., a realistic depth perception, with reduced image deformation too.

Research Question: What is the correct headset display FoV?

Figure 62: Conceptual representation of observation through HMD showing the FOV and the observed sphere.

Implementation (Zoom effect):

The developed technique consisted of mapping the acquired image into a sphere using a prefabricated Shader in *Unity* called *Skybox*, as previously introduced within section 6.2.1.2.

To obtain a zoom effect one wishes to act on the virtual camera position, implemented on Unity software by the *CameraRig* function. However, the related camera game-object does not allow for changing the FOV. Our proposed solution to achieve the zoom effect is based on moving the virtual camera position (keeping the FOV fixed) away from the center of the sphere. This results in the same effect as when you change the FOV. Figure 63 illustrates the proposed solution.

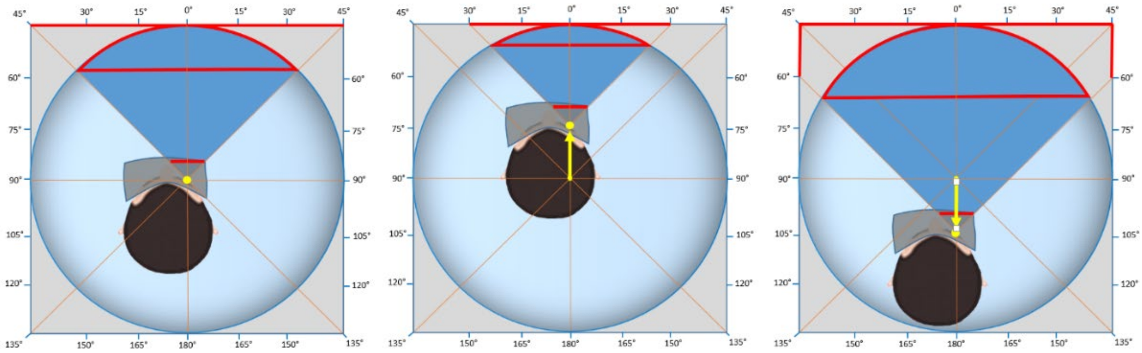
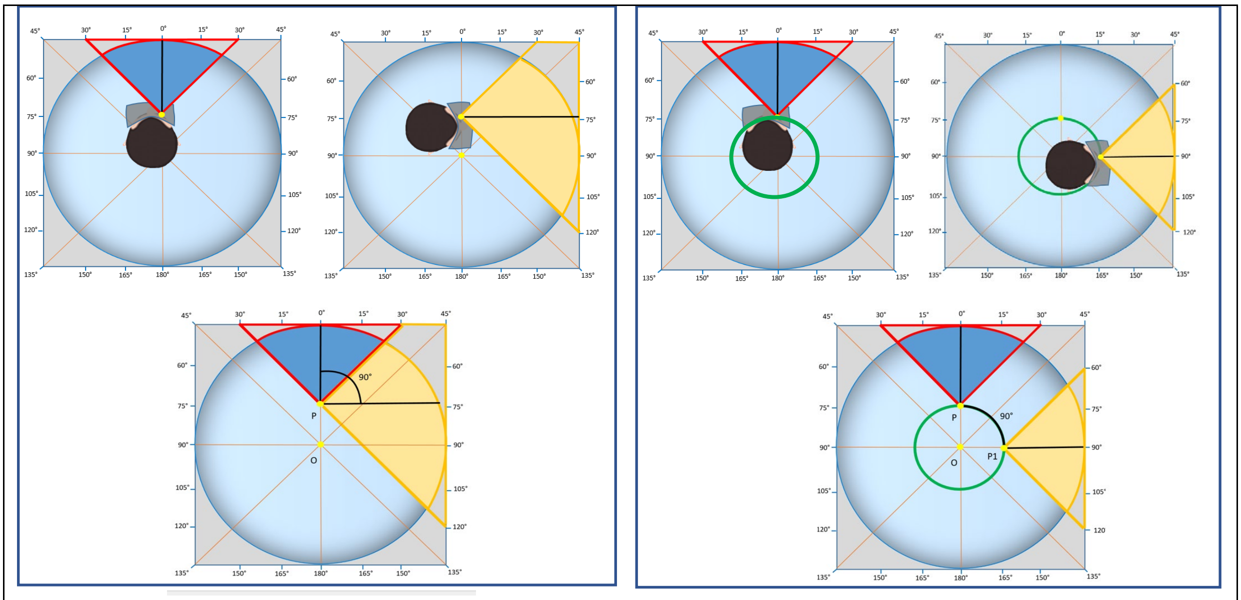


Figure 63 – The proposed solution for recreating the zoom effect: observations with three different positions.

The zoon effect works, without distortion, only when the observer is perpendicular to the surface of the sphere. Figure 64(a) shows that as the observer rotates his head, the FOV also changes creating distortions in the perception of distances.

To extend the proposed zoom effect solution to any user movement (within the texture-mapped spherical environment), a new method was devised. We called it: *Concentric Spere*.

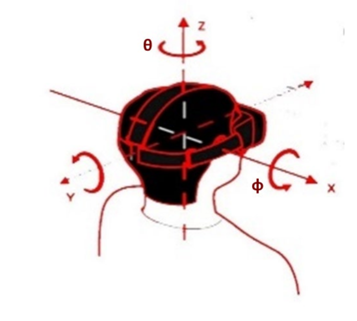
Figure 64 (b) shows that, applying the Concentric Spere method, when the observer rotates the head, the observation position also changes, keeping the FOV fixed and the view perpendicular to the surface of the sphere. Therefore, by applying this method, it is possible to vary the FOV without creating distortions on the observed image.

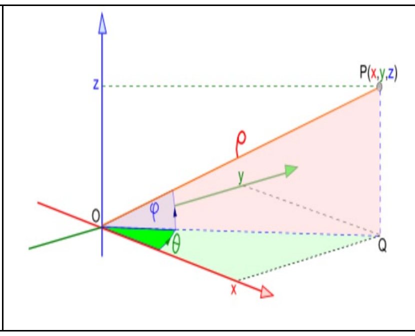


(a) *(b)*

Figure 64: (a) Rotation without Concentric Sphere method; (b) Rotation with proposed Concentric Sphere method.

The position of the observer is computed by applying the formula shown in the right-hand side of the figure 65. The formula shows spherical coordinates. Figure 66 shows examples that include a full head rotation.





$$\begin{cases} x = \rho\cos(\varphi)\cos(\theta) \\ y = \rho\cos(\varphi)\sin(\theta) \\ z = \rho\sin(\varphi) \end{cases}$$

ρ = distance from the center of the sphere

θ = Rotation along the z axis

φ = Rotation along the x axis

Figure 65: The proposed Concentric Sphere method: graphical and mathematical explanations.

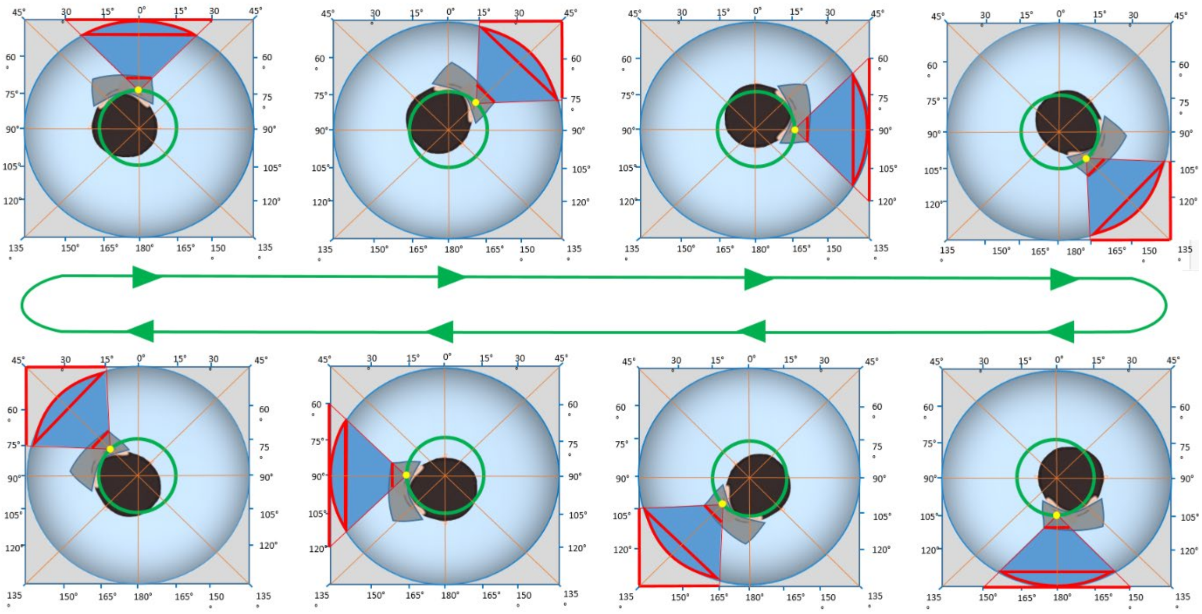


Figure 66: Examples that include a full head rotation using the proposed Concentric Spheres method.

The proposed Concentric Sphere method was tested through a number of trials. The evaluation was conducted looking for distortions along 20 different viewing paths. The outcome showed a clear improvement when navigating within the sphere with Concentric Sphere method compared to a system that did not include this method.

6.3 Assessing HMD Depth Underestimation

This part of the PhD project involved research, development and assessment on realistic depth perception. The work is set to start with assessing and therefore verifying the HMD typical underestimation of perceived distances, to then proceed in the next section with providing insight and a procedure about how to improve HMD depth perception.

6.3.1 Testing Environments

The use of two different environments was proposed for the experiments to test for distortions and misalignments between real and virtual views. The following environments were chosen:

1) *Corridor*. This was the corridor area in front the UH VR & Robotics laboratory. It is a long rectangular area with size 3.30 x 30 meters, chosen because it allowed for testing long distances to viewer. Figure 67.a shows the environment.

2) *Room*. This was the UH VR & Robotics laboratory main room. It is a rectangular area with size 5 x 6.5 meters, chosen because it allowed for testing in a space that was wider. Furthermore, we had more control in terms of reserving the space and arranged as needed. Figure 67.b shows the environment.

(a) *(b)*

Figure 67 – Corridor environment Room environment

In order to assess how realistic is the user's depth perception, we measure the error in user's distance estimates. We ran a number of test trials where distances to specific objects were estimated by users. Those measurements were acquired using the blind-walking method. This is one of the most popular methods used in the literature. The blind walking technique is described in [23]. Users are instructed to look at the target location, then close their eyes and reach for it. The distance travelled is then measured.

6.3.2 Photo-Based Omnidirectional Rendering.

The method used for creating our photo-based virtual environments followed the method introduced in section 2.1. A 3D-360-degree photograph was acquired and then mapped onto a sphere. The sphere and its photographic texture represented the 3D model used by the computer-graphic engine to render in real-time a S3D view of the represented scenery.

The user observation viewpoint was by default set at the spere's center. As for the other viewing elements, they followed the guideline underlined in the proposed photo-based VR system concept (see section 5.2 for details). The S3D effect of the rendered image, generated by the software, had an IOD similar to that average human interpupillary distance (IPD). This is approx. 6.5 cm. The FOV, was chosen such that it would deliver the same impression of the human vision FOV.

- Photo Acquisition. This was done in line with what presented in section 5.4.5, but we used a different camera. We captured the scene with our high-resolution three-dimensional 360-degree camera system (Insta360 Pro, [10]). Figures 68.a and 68.c show a moment during the acquisition of environment photos.

- Each environment needed to be calibrated and prepared for testing. This included putting a number of reference lines on the floor and signposts on the walls. Those references were needed to support calibration, to make images misalignments more clearly visible, and to allow for measurement of the presented displacements. Through the use of the Insta360Pro software it was possible to see in real time what was being filmed by the camera. Figures 68.b and 68.d show an example view of the two environments during the calibration process. The calibration along the horizontal axis was performed with the support of the software named: Insta360 Pro Camera Control App [167]. The red line needs to align with the wall signposts.

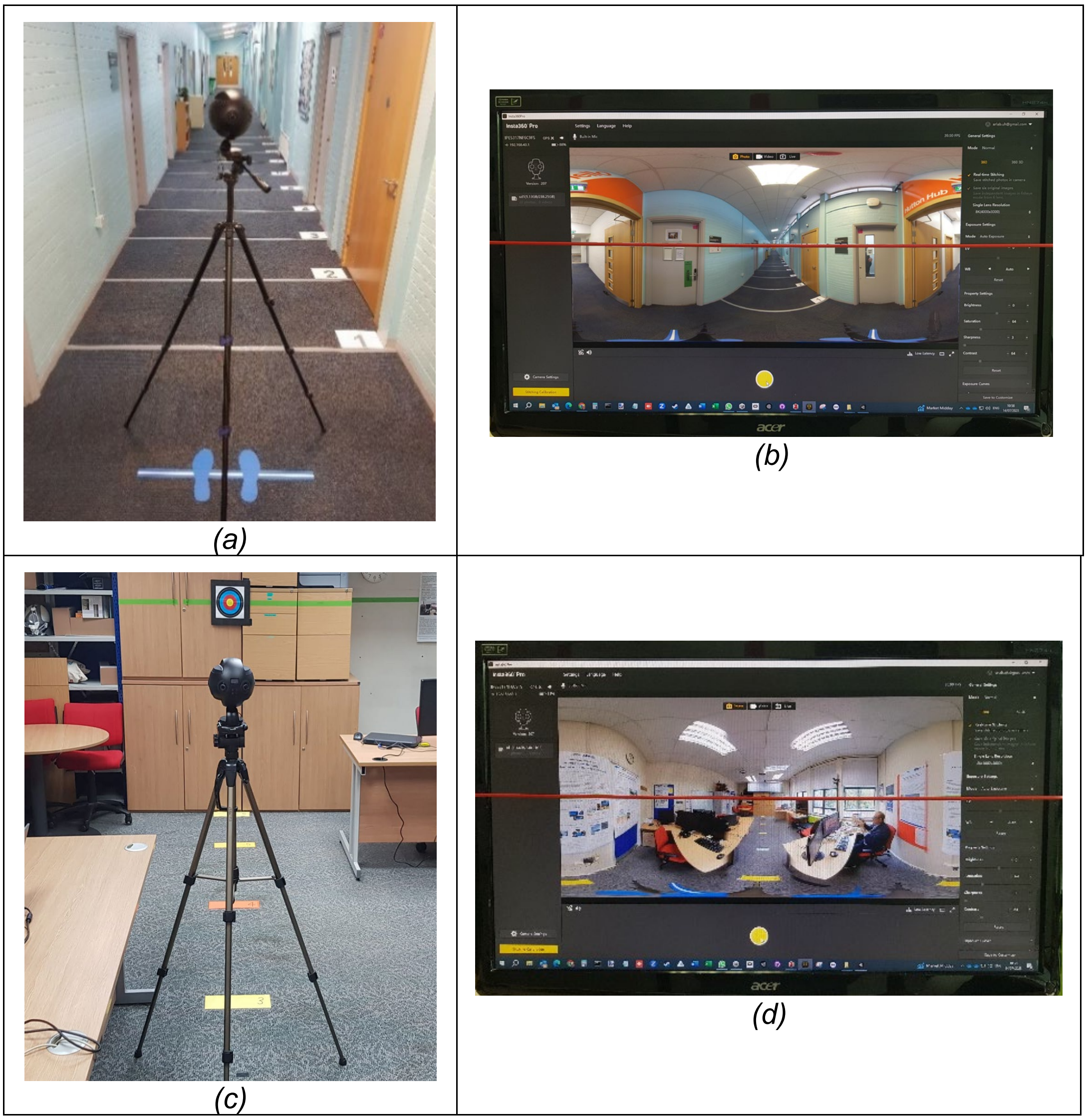


Figure 68 – Left: Environment acquisition, (a) Corridor (c) Room. Right: Monitoring and Calibration, Insta360 Pro Software.

- Photo Processing: Figures 69.a and figures 69.c show the acquired stereoscopic image (left and right) of the two environments captured with the Insta360 Pro camera. Figures 69.b and 69.c show the acquired stereoscopic image being mapped into a spherical object by the Unity Software [168].

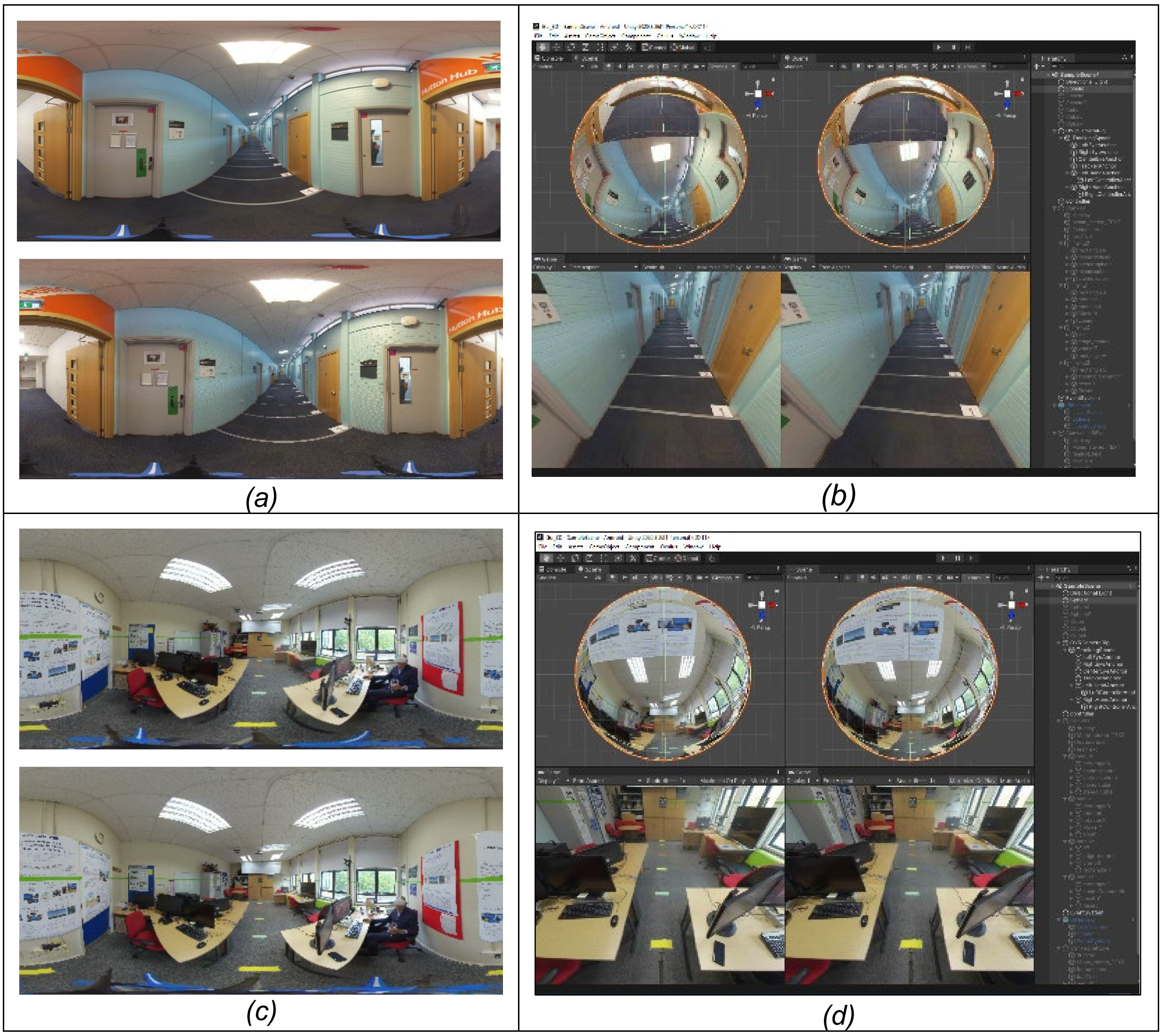


Figure 69 - (a) and (b) The acquired stereoscopic image representing the corridor environment and the room environment, respectively. (b) and (d) The acquired stereoscopic image being mapped into a spherical object by the Unity Software.

- Photo Visualization. The figures 70.a and 70.b, 71.a and 71.b, show the two testing environments and the user during experimentation. The figure 70.c and 71.c show an example of the rendered environment image for the corridor and room environment, respectively. The rendered image is generated by the Unity software from the spherical 3D model. We call the rendered image the virtual view.

6.3.3 Exploiting the Passthrough Option

A useful feature that latest VR headsets provides is commonly called *passthrough*. This allowed users to see the real world surrounding the HMD through the cameras the HMD is equipped with. This option was exploited to help understand and design our assessment, and to help with setting up test-users' views.

The test assistance who guided the experiment (the test monitor [49]) also used the *passthrough* option to follow user's actions in a simple and clear way. The test assistant could switch among real, virtual and mixed reality views, and get immediate visual feedback while running the experiment.

The figures 70 and 71, right-side images show three different example views of the two environments, Corridor and Room, where the *passthrough* option was used to represent the real and the mixed views. The images show S3D views.

They are:

- *(c)* *Virtual view*. Photo-based computer-generated image of the corridor.
- *(d)* *Real view*. Live image of the corridor captured through HMD's cameras (passthrough only).
- *(e)* *Mixed view*. Overlapping Real and Virtual views of the corridor (both passthrough and virtual view).

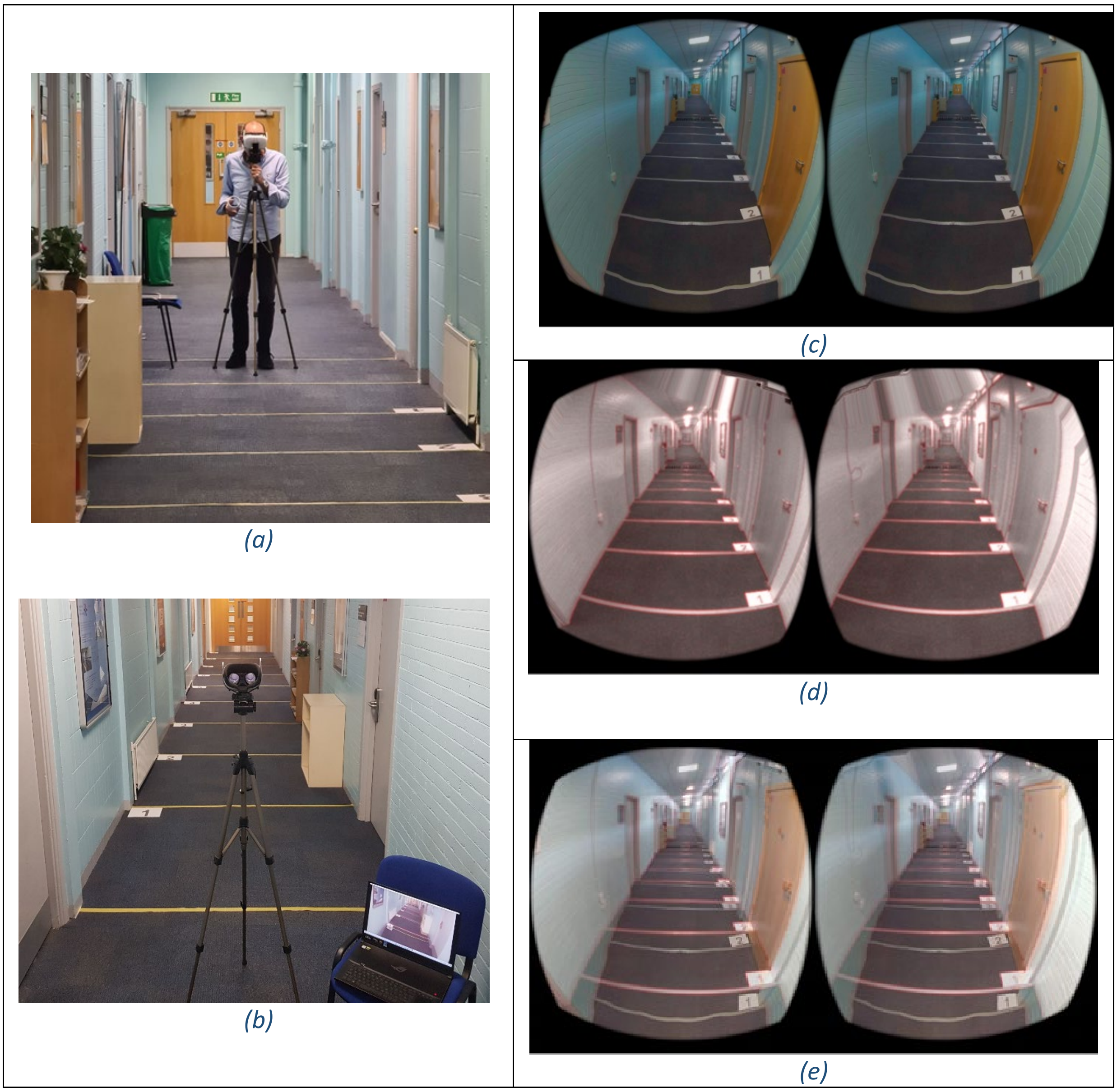


Figure 70 - (a) and (b) HMD Visualization; (c) Virtual Environment; (d) Real Environment with passthrough; (e) Mix Reality: Virtual and Real Environments together.

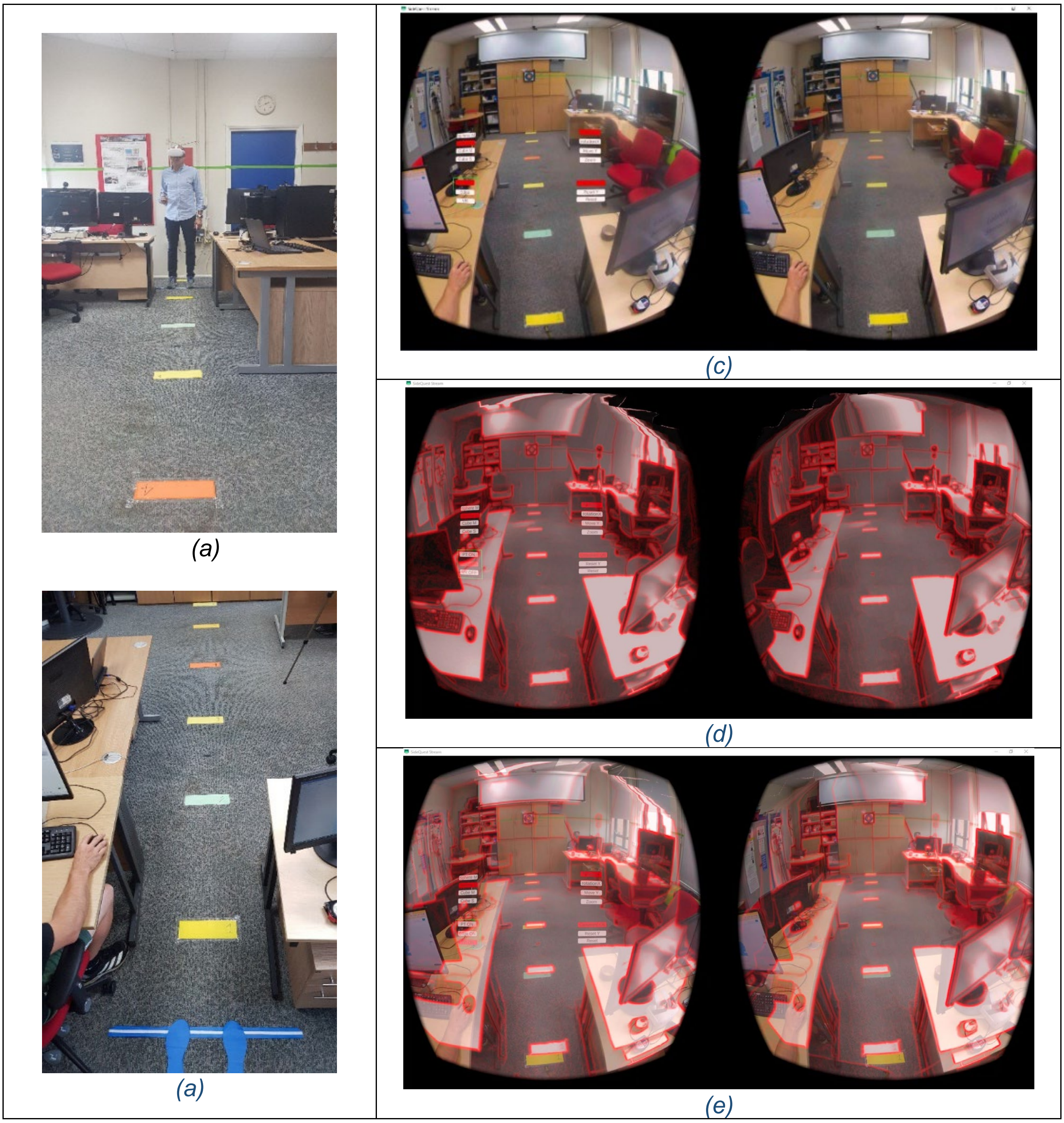


Figure 71 - (a) and (b) HMD Visualization; (c) Virtual Environment; (d) Real Environment with passthrough; (e) Mix Reality - Virtual and Real Environments together.

6.3.4 Experiment Execution and Outcome

The assessment started by performing several observations of the same environment area represented as virtual view or real view. The user's perceived distance to objects where measured, which resulted to a noticeable estimate displacement. This was made even more noticeable to test assistant by outlining in red color any detected edge in the real image.

The two environment images illustrated in figure 70.e and 71.e, clearly show the views misalignments. Interestingly, the misalignment varied according to objects distance, which was highlighted by discrepancies in the floor lines set in the corridor environment.

We ran a number of test trials to estimate distances to specific objects following the above-mentioned blind walking technique. The first thing that was clearly noted was that there was an area of the environment where no displacement appeared between virtual and real views of the same object. This was clearly shown by the floor lines in the virtual and real images overlapping each other. We call this area *no-displacement zone*.

If we observed the environment space between the no-displacement zone and the viewer's position (we call it the *preceding area*), we could clearly see that the object displacement was growing towards the viewer. In particular, the object in the virtual view clearly appeared to increase in size compared to the object in the real view. The increase grew as it got closer to the viewer.

Because of the perspective effect, the above outcome had the consequence that objects in the virtual view located in the *preceding area* (i.e., between the no-displacement zone and the viewer), appeared closer to the viewer when compared to the same objects in the real view. This led to a clear users' underestimation of distances when observing the environments virtual views. This result was in line with most of the related literature works [23],[152], [169]. We could not find clear reasons leading to the measured underestimation of distance, as also typically stated in the literature.

In the part of the *preceding area* closest to the viewer, a double image effect could often be seen due to the excessive parallax of the generated stereoscopic image. It had been indicated that this effect clearly led to the perception of a larger size of the observed object, which would then add to the previous effect, and further contributes to the underestimation of the distance in the observed objects. The observed shape of the object also played a role in this area. Figure 71.e, shows this effect.

Interestingly, we could also observe the object displacement between real and virtual views grew in the reverse order when moving away from the no-displacement zone towards a direction running opposite to viewer.

What described so far, indicated the relevant role played by the no-displacement zone in terms of images misalignment and accuracy of distance perception. This is a new relevant finding that we do not find in the related literature. We conclude that the literature works assessing distance underestimation in VR, had likely run experiments across the *preceding area* only.

Overall, the outcome of the presented evaluation can be summarized in the following points:

a) *No-Displacement Zone*. There was an area of the environment where no displacement appeared between virtual and real views of the same object.
b) *Increasing Displacement*. Out of the No-Displacement Zone, the environment areas showed clear and varying displacements. The displacements increased stepping away from the no-displacement area.
c) Object displacement between virtual and real views increases in the *preceding area* the farer the object is from the no-displacement zone. In the *preceding area* the objects in the virtual view appear larger than in the real view.

d) Object displacement between virtual and real views increases when moving away from the no-displacement zone towards a direction opposite to viewer. In this area objects in the virtual view appears smaller than in the real view.

e) A double-image effect may take place in a part of the *preceding area* closest to the viewer. This may result in a further increase of the observed objects appearance in the virtual view, which can contribute to distance underestimation.

f) *Camera Pose*. Making change in camera position (both along the horizontal and vertical lines) or in camera orientation, resulted in changes in the no-displacement area.

g) *Viewer's Pose*. Making changes in viewer's position and viewing orientation resulted in changes in the no-displacement area. This happened even in case of minor changes (few centimeters).

h) *Software COP*. Making changes in the software-camera center-of-projection (COP), which is managed by the rendering software, resulted in changes in the no-displacement area.

i) User's distance underestimation is confirmed when observing a virtual view of the represented environment within the environment portion defined as *preceding area*.

6.4 Improving HMD Depth Perception

This last part of the PhD project involved research that exploited previous assessments and development on realistic depth perception to devise solutions to reduce errors in perceived distances underestimation.

An investigation on VR portrayed photo-based realistic viewing involves photographic image acquisition, processing and visualization. It can therefore be easily deduced that this investigation can become very complex due to the numerous system parameters involved. These parameters can, for example, be those relating to the type of optical system of cameras and viewers, acquisition and display settings, those relating to software processing and graphic rendering, image corrections, view-porting projection, etc. Clearly, an exhaustive investigation is realistically impossible, and we need to greatly delimit the parameters we want to focus on, based on available resources.

In addition, there is a degree of difficulty in isolating specific system elements of VR systems, which becomes even more difficult when one targets practical applications, and therefore wish/need to use off-the-shelf ("inflexible") hardware.

A major question is then about whether we can still do something to improve depth perception in photo-based VR representations observed through an HMD display. This question was definitely asked, and the answer was that we can still have meaningful results by both relying on same literature outcome and by focusing on few system elements expected to play a key role.

The literature indicates that an important role is played by the setting of the display view. This has therefore become an element on which we wanted to focus. Also, with our display images being photo-based, we also had to consider the image capture setup.

In our system we had the following IOD related settings: (1) the stereo-camera acquisition system had got a camera baseline (i.e., the stereo-cameras distance) equals to the average human inter-ocular distance (IOD = 6.5 cm); (2) the two virtual cameras of the visualization software (*Unity3D*) were set to also have their baseline equals to the average human IOD; (3) the display system of our HMD (*Meta Quest 2*) allowed for setting the headset IOD according to user's interpupillary distance (IPD).

The above indicates that our system followed the indications of the proposed system concept (section 5.2) to minimize distortions that occur due to a mismatch of hyper or hypo stereo settings. Unfortunately, distortions cannot be completely overruled because this would require setting up the software and more importantly the settings of the capture camera, according to each individual user. Such item settings cannot be obtained both for practical reasons and because we are using off-the-shelf acquisition and processing systems which do not allow such tuning. Using off-the-shelf systems is part of our goal, so we don't want to change that.

The good news is that our visual brain can mitigate or even eliminate misalignments in IOD if these are contained to millimeters or even few centimeters (depending on the distance to the objects we are observing). This has been confirmed in many literature works.

The user study 5 described in section 6.1 also confirmed that containing image distortion by following the indications of the proposed system concept could also be helpful towards providing a realistic distance perception. The concept of aligning the IOD "pipeline" seemed to work as it led to an error in distance estimation in line with the literature (below average in some points).

6.4.1 Determining The No-Displacement Zone Location

A pilot investigation was decided to be conducted, which would focus on discovering specific elements that would play a relevant role in depth perception beside those related to the camera-environment geometry considered in user study 5.

The starting point was the outcome of the evaluation described in section 6.3.1. This led us reflect on about how to counterbalance the underestimation of distances that users perceived. We wanted then to check what caused the underestimation by testing object misalignment for a greater number of images taken by a set of different positions inside in our laboratory environment. We looked for No-Displacement Zone and Increasing Displacement.

Thirty different 360 images of our lab-environment were captured from a set of positions across two horizontal and two vertical lines. The figure CC shows the positions. The results analysis was then facilitated by the use of the *passthrough* option, which allowed for overlapping the virtual and real views.

The two settings that appeared to clearly play a role were:

a) Viewer's height.
b) Camera horizontal orientation.

They clearly affected the alignment of objects in the two virtual and real views. In particular, they determined the position of the No-displacement zone. As for the Increased Displacement, it appeared to follow the behavior described in the outcome of the evaluation in section 6.3.1.

In summary, clear evidence was gathered that the height of the viewer and the horizontal orientation of the camera have an effect on the position of the no-displacement zone and increased displacement, leading to a different depth impression and distance perception for each individual user.

The above considerations indicated the need for calibrating acquisition and visualization setups to correctly set the camera horizontal orientation, and to make acquisition according to the individual user.

A two-step procedure was then proposed following the actions below described:

a) The height from which images are acquired should be equal to the height from which images are viewed.

b) The orientation of the camera must be precisely calibrated with respect to the horizontal plane and this orientation must be replicated on the displayed image.

6.4.2 Experiment Design and Execution.

An evaluation of the proposed two-step procedure was proposed, where the camera position was set according to user's height and the camera orientation was horizontally calibrated.

An evaluation was then conducted with consisted of 30 test trials to estimate distances to specific objects following the above-mentioned blind walking technique.

The outcome of the performed trials outlined the following facts:

- The accuracy of distance estimates was, as expected, improved by the proposed acquisition and display method. The most accurate distance estimates were recorded in the area around the No-Displacement Zone.
- A degradation of estimates was noted in areas from the no-displacement zone.
- A degradation of estimates was noted when images were acquired at a different height from the user's height.

The estimated distances when observing the virtual view still showed a general underestimation. Nonetheless, the results were promising because they showed higher accuracy compared with previous trials where the proposed method was not applied.

The above outcome provided a clear indication on a direction towards achieving a more realistic depth perception, which we deem represents a relevant outcome of our research and a direction that is worth to further investigate.

6.4.3 Viewpoint Elevation and the Viewpoint Retreat

The main drawback of the method proposed above is the time-consuming acquisition and calibration processes. They needed to be repeated for each different user's height, so that users with different heights can realistically perceive depth.

To reduce the number of acquisitions at different heights, a solution would be to choose a subset of user's heights and exploit them as reference for any user's height, e.g. through using the images acquired from the closest position. This solution may still require a good number of height acquisitions to produce a realistic depth estimation.

Alternatively, one could think of having the acquired 360 environment image mapped on sphere by a chosen viewpoint set for example at an established user's height, and then during observation move the visualization viewpoint along the vertical axis according to observer's height. We call this method: *Viewpoint Elevation*.

A different approach to reduce distance underestimation in virtual views is proposed to counterbalance the underestimation by shifting the observation viewpoint backwards on the x-axis. The observation viewpoint is the center-of-projection of the image-rendering process, typically set at the center of the mapped sphere, which is controlled by software. The amount of such observation viewpoint shift needs to be estimated. We call this method: *Viewpoint Retreat*.

Both the *Viewpoint Elevation* and the *Viewpoint Retreat* approaches were tested by analysing their effect in terms of misalignment between virtual and real views. The camera views were calibrated following the produce described in figure 72. The viewers movements during the observation included the use of the *Concentric Circle* technique described in section 6.2.

A specific software interface, figure 73, was designed to allow for the adjustment of several related settings which included the *passthrough* option for seeing the resulting real and virtual views simultaneously. The HMD observed views were also replicated on a desktop monitor at the disposal of the test assistant.

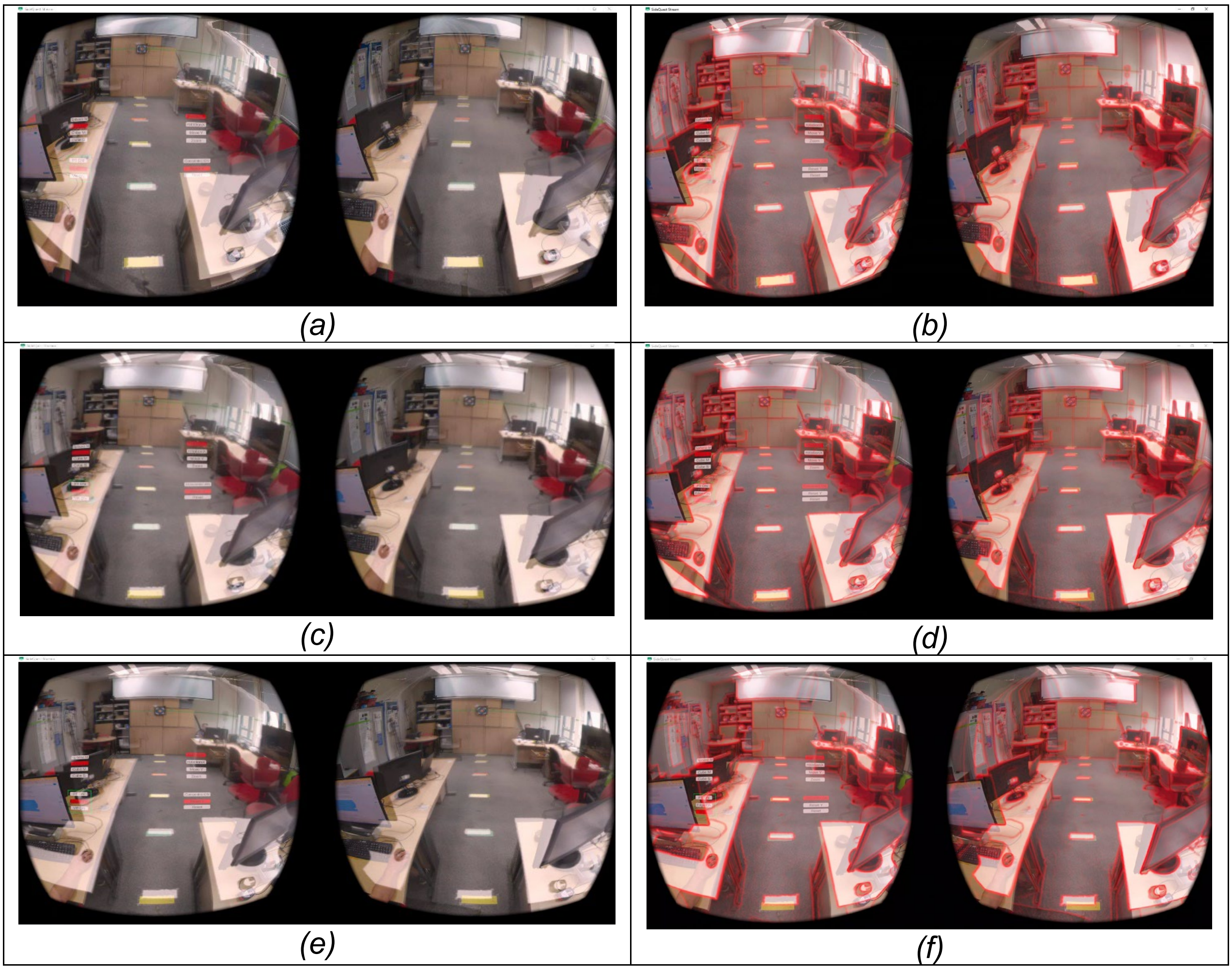


Figure 72 - All images are mixed reality. Landmarks are the rectangles equally spaced from the bottom center of the images to the back of the room. The closest landmark to the bottom center is landmark 1. Images (a and b) are aligned to closest landmark, images (c and d) are aligned to next closest landmark, images (e and f) are aligned to landmark 3. The images in the right column have a red tint that highlights the point in focus.

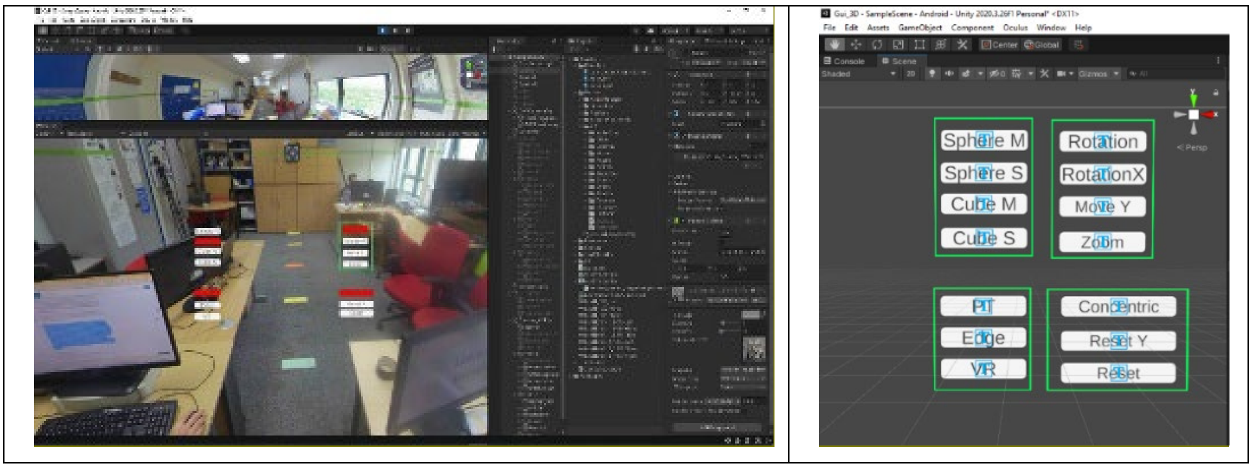


Figure 73 – The left-hand side image shows the testing environment as seen from the headset. This image also shows the developed specific software interface. The right-hand image shows the list of software options. They include: Sphere mapping (mono, stereo), cube mapping (mono, stereo), rotation (x-axis, y-axis), move (y-axis), zoom (concentric spheres), passthrough, passthrough with enhanced edges (red colour), virtual, reset positions (3 options).

Applying the *Viewpoint Elevation* to adapt the viewpoint position to user's height, leads to a noticeable image deformation during observation between the higher and lower hemispheres. This is certainly a drawback of the method because it affects objects' perceived shape and their proportions. On the other hand, the location of the No-Displacement Zone appears backwards, which has the effect to render objects in the virtual views further away, promising to reduce distance underestimation. The performed pilot test measurements gave contrasting results, which called for further test trials to get more precise conclusions.

Applying the *Viewpoint Retreat* to reduce distance underestimation leads to a noticeable reduction of object size, which is reflected on higher distance estimates. The No-Displacement Zone is moved further away from the viewer promising to provide higher accuracy on user's distance estimates, which could clearly be seen in our pilot test measurements. The retreated viewpoint affects the observation FOV as it increases the environment observable area. This effect did not seem to have much affected users' estimates. It is expected this depend on the shift amount. Further test trials would be needed to understand how this parameter actually affects depth perception.

The results of the pilot testing were encouraging for Viewpoint Retreat. The reason is justified by such shift backwards not producing deformations on the horizontal place. Rather, some deformation occurs when the camera rotates around the y-axis.

6.4.4 Adjusting the No-Displacement Zone

A clear outcome that emerges from the tests conducted is the sensitivity of the No-Displacement Zone to changes in observation viewpoint. In some cases, the changes were minimal, while in other cases they were greater. Above all, the effect on the size of the object and the consequence in terms of accuracy of the estimate seemed clear. Image deformation emerged as a different issue. It affected objects appearance deformation rather than their size.

A further assessment had therefore been proposed in which the location of the No-Displacement Zone is changed by a rotating the sphere (on which the image is mapped) along the x-axis.

Since this rotation is along the x (horizontal) axis, it clearly refers to horizontal calibration. However, to represent an accurate horizontal calibration, we need the x-axis to be positioned at the user's height.

Interestingly, a rotation of the sphere along the x-axis changes the alignment between virtual and real images, moving the No-Displacement Zone in one direction towards or away from the viewer. Clearly, this should affect the distance estimation because it changes the size of the object's appearance. We therefore wondered whether such rotation along the x-axis might be acceptable to serve as a horizontal calibration. If acceptable, it would speed up the calibration process because in this way horizontal calibration can be performed regardless of the individual user's height.

The *Horizontal Calibration by Sphere Rotation* was therefore proposed for testing. This method was proposed to be coupled to the previously described *Viewpoint Retreat*. In this way we could potential achieve an efficient and effective method to counterbalance the user's distance underestimation emerging in virtual view representations.

The setup procedure for virtual view alignment follows the photo-based graphic rendering through spherical mapping. The No-Displacement Zone represents the reference target. This should represent the environment area where we wish to achieve the best performance in terms of realistic depth estimation.

The proposed testing procedure is then as describe below:

1. The user estimates the distance to object *obj1* by direct observation. The user then closes his/her eyes and walks towards *obj1* until he/she reaches it.
2. The user estimates the distance to object *obj1* by indirect observation (using the passthrough option). The user then closes his/her eyes and walks towards *obj1* until he/she reaches it.
3. The user estimates the distance to object *obj1* by virtual view observation. The user then closes his/her eyes and walks towards *obj1* until he/she reaches it.

The steps 1, 2, 3 should be repeated for each chosen set of objects (*obj1*, *obj2*, …*objN*). The procedure must follow a counterbalanced schedule regarding the 3 different views.

Works in the related literature have indicated a typical error on direct observations equivalent to approximately 8% of the actual distance; while the error on indirect observations was equivalent to 20% of the actual distance; and finally the virtual observation error was equivalent to approximately 26% of the actual distance, [170], [23].

The outcome of 30 estimation trials carried out with 3 objects was very encouraging. It showed an average error of 16%.

6.5 Conclusions

It seems impossible to identify a unique procedure to solve the problem of realistic depth perception in VR portrayed environments. There are many parameters involved across the different processes. In case of photo-based VR rendering, those processes and parameters must include the acquisition process in addition to the visualization and observation processes.

Chapter 6 presented a study into distance underestimation in photo-based VR, which aimed to identify possible ways to address distance underestimation that is typical of virtual view observations through VR displays. The presented studies included a user study (section 6.1), a number of focused developments (section 6.2), a target assessment (section 6.3) and a method to reduce distance estimation underestimation, by counterbalancing misalignments between real and virtual views of the same observed environment (section 6.4). The method proposed in section 6.4 represents the base for a future full evaluation.

7. Conclusions and Future Research

The ultimate objective of the PhD project was to study and assess what we called photo-based VR, in order to understand its potential in light of the recent spread in the use of high-resolution 360 cameras. It was relevant to understand whether photo-based VR could represent a new and popular way to create VR observations, typically representing remote locations. Being such solution expected for observing real places, realistic depth perception was expected to be a key aspect to investigate. This would be particularly relevant in applications related to the remote monitoring and observation of naturalistic and touristic sites, and in realistic scene reconstruction useful for surveillance systems, control panels, and operator's training.

To investigate the potential of photo-based VR and to assess its performance, a need was identified to evaluate the role of immersion and graphical visual experience. Therefore, the research started with assessing what advantages immersive visual experience can already provide, to then develop and assess photo-based VR. Finally, a focused investigation addressed what we called true-dimensional visualization, i.e., a visual observation that can provide correct perception of space dimension and objects size.

7.1 Summary and Contributions

This thesis presented the work done in the PhD project through its 7 chapters.

Chapter 1 introduced the project, provided an overview on immersive visual experience, and on outlined the proposed investigation.

Chapter 2 provided the reader with the background knowledge most needed to comprehend the work presented in the following chapters.

Chapter 3 presented the idea and motivation for the proposed investigation, and the associated research development plan. The investigation consisted of three main parts, which respectively focused on assessing the immersion advantage, the photograph advantage in VR, and the true-dimension advantage (defined as realistic depth perception).

In *chapter 4*, the provided sense of immersion and its related advantages were assessed within the user studies 1 and 2. The first study addressed pilot training (including flight deck training and skilful drone teleoperation). The results indicated that the use of an immersive system such as the HMD led to a more effective training in terms of precision in all tasks and in terms of completion-time in case of S&F tasks, scoring well also in terms of isolation with consequences in terms of presence and depth impression. The desktop monitor proved to be suitable for simple and repetitive actions such as operating the check-list at deck, and for driving the drone into a simple narrow tunnel. It also scored well on comfort. The overall conclusion was that technology providing immersion is ready to be widely adopted in pilot training, where its superiority against the DM appears relevant. The second study addressed user-scene interaction with two different interaction modalities for viewpoint change (HMD-J using joystick, HMD-H using head rotation). The HMD-H had a clear advantage in terms of response time and naturalness of operation, whereas both modalities achieved the same operation accuracy.

In *chapter 5* an assessment on the use of photographs as 3D model for VR applications (photo-based VR) from which the graphical rendering could be generated, represented the common base for the user studies 3 and 4. The first study addressed the role played by screen and location focusing on the role played by pixel-density and location's illumination and represented object distances. The results indicated the contribution of higher *pixel-density* was positively felt and it prevailed over better lighting specs, leading to a significant improvement in some of the *visual realism* and *presence* factors. Image lighting factors were felt to contribute to realistic viewing. The locations' environment showed to affect *spatial presence* and *involvement* eliciting some *emotions* too (*enjoyment, relaxation* and *anxiety*). Both location's illumination and distance to objects contributed towards *depth impression* (respectively in the island and cave environments). The location generally played a stronger role than the display, which proved that good quality displays are "transparent" to scene content. Interestingly, the display characteristics were still able to further enhance *image-lighting*, *spatial presence* and some elements of *visual realism* and *emotions*. The second study focused on the role played by the sense of place provided by previous knowledge. The results indicated the effectiveness of immersive technology combined with photo-based VR as users knowing well the observed place scored high *visual realism* and *presence*. Interestingly, the outcome for place-familiar users proved that previous knowledge can positively enhance the perceived *realism* and *presence*, e.g., by bringing memories. This was a fascinating aspect worth future investigation.

The performed studies proved the overall effectiveness of the visual experience provided by 3-D omnidirectional photorealistic images, observed through a VR-headset, and visualized according to the proposed system concept. We deem the outcome of both experiments was particularly positive if we consider the limitations in terms of static images and a choice for system elements constrained by the use of off-the-shelf devices (including camera and visualization systems).

In *chapter 6* the possibility to provide a user with a realistic perception of space dimension and objects size was investigated within photo-based VR. The first step was conducting the user study 5 to gain an indication on the effectiveness of our photo-based VR rendering and the followed system concept, in terms of depth perception accuracy. The results showed the depth sensation was undoubtedly supported by environment illumination and distance to objects, e.g. affecting *anxiety* in the cave environment. Distance estimation accuracy was in line with that indicated in the literature (better in average in some points) which show the effectiveness of our system settings.

A series of focused assessments were then conducted to understand the role of some relevant parameters, above all the FOV. These assessments have sometimes been followed by some important ad hoc developments. Most notable was the Concentric Sphere method, which has been shown to effectively remove zoom distortion when navigating.

A further step of investigation was conducting an assessment on the causes of distance underestimation typical of virtual view observations through VR displays. a number of assessments were run, also including ad-hoc developments. A specific procedure was then proposed to reduce distance underestimation, by counterbalancing misalignments between real and virtual views of the same observed environment. The results from a number of pilot trials showed the proposed procedure to be very promising and suggested a full-scale user-based assessment as future work.

7.2 Lesson Learnt and Future Directions

7.2.1 Lesson Learnt

The outcome of the project and the lesson learnt clearly indicate that photo-based VR represents a valid alternative to traditional visual rendering, which can successfully be proposed on applications that require an efficient and effective way to remotely render real spaces.

The acquisition of wide-angle/omnidirectional images can be fast and economical. Capturing 3D images as well as visualizing images following specific settings, appear critical in order to achieve a realistic reproduction of depth.

We have learnt that immersion can be really advantageous in applications that require accuracy, precision and naturalness of operation, whereas a non-immersive setting can be a good complement to immersive technology, which provide more comfort.

We have learnt that photo-based VR is more than traditional texture mapping and can effectively represent virtual and remote environments.

We have learnt that photo-based VR has unique added values. Being completely surrounded by photographic texture (as it is the case in photo-based VR) can convey much stronger sensations than traditional visual rendering. Image-lighting can more strongly affect visual appearance and emotions such as relaxation and enjoyment, but also fear, can more easily be communicated.

We have learnt that realistic depth perception is challenging and that accurate settings may be impractical, but we have also learnt that specific remedies are those investigated in the presented research, can be at hand and be very effective too.

7.2.2 Potential Applications

Photo-based virtual reality is widely applied due to the widespread use of 360° cameras. Omnidirectional image acquisition is available to most of the population due to its low cost. Industries and commercial businesses can also easily take advantage of this technology and target even the most advanced and capable version of this technology, such as those systems providing high resolution (e.g. 8k) and stereoscopic 3D capture too.

Photo-based VR can also rely on computer generated images, which may lack to provide highly realistic photometric effects but can on the other hand allow for any scenario and perspective view to be generated with high image quality.

Graphical mapping of omnidirectional images (either captured with cameras or computer-generated) would then be the next processing step to go through in order to render scene views from current user’s perspective. This operation can rely on simple 3D models such as spheres or cubes, which despite their relative simplicity, are still capable of significantly influencing the user's sensations because of the rich graphic texture. This can even be the case with texture-mapping onto simple 3D models, which represents a limited computational effort for today's rendering engines. Therefore, this is convenient and quick to accomplish process.

Application sectors such as architecture, urban planning and construction can greatly benefit from the proposed system. Likewise, applications that provide virtual visits to places. Design tasks such as virtual prototyping of objects and vehicles can also benefit from photo-based VR technologies. The proposed approach to VR visualization is useful both for observation and design activities, as well as training. All these activities accelerate the knowledge of the target application thanks to the high level of realism provided.

To maximize the experience, viewing should be via a VR headset. This system provides a great sense of isolation from the surrounding world, which combined with the wide/360 field of view improves presence and operational performance.

The additional possibility of having a true-dimensional visualization on photo-based VR representations, further extends the application possibilities. This happens because a realistic visual appearance is combined with a realistic perception of the spatial dimension. This is of particular interest when reproducing indoor environments and when the realistic impression of depth and distance estimation is a relevant feature.

For example, when looking at a remote location that one wants to explore for the purpose of buying or renting, it is critical to understand its actual spatial dimensions. This advantage can be extended to all applications that would benefit from realistic space observations. Visually understanding the true spatial dimension is also important for teleoperation training, including that of robots or medical instruments such as endoscopes. Pilot training actions can also benefit greatly from true-dimensional visualization. This ability allows pilots to perceive realistic dimensions of cockpit panels, which accelerates their knowledge of real dashboards and helps them understand the real distance to objects and their real size.

Training for mechanical operations, including assembly, can also be significantly accelerated by photo-based virtual reality with true dimensional display.

7.2.3 Future Directions

Further research would be required to further investigate and assess both photo-based rendering and true-dimensional visualization. This research can also be oriented to specific applications. Photo-based VR and true-dimensional visualization can also be researched separately, as they can also be applied to VR individually and independently.

A possible future research activity can include conducting a formal study of the approach presented in section 6.4. For this purpose, the strategy and topic of investigation has been presented, which can represent the base to use for designing a user study.

A direction of future investigation can be including dynamic scenes in photo-based VR through the use of 360-degree videos or live streaming and enhance user's interaction allowing for greater user's movements such as teleport.

Linking immersive visualization and photo-based visual realism with studied addressing psychophysical aspects would also represent a future direction of investigation. This could be addressing specific applications, such as rehabilitation therapy and operational trainings.

Concerning potential directions towards achieving a more automated process for true dimensional visualization, the following observation could be considered.

The calibration of the observed space for real dimensional visualization in photo-based VR representations can be performed through an ad hoc technique, such as the one proposed and tested in the thesis. This simply consisted of carrying out two operations aimed at redefining the no-displacement zone. The two operations have been named as: *viewpoint elevation* and *viewpoint retreat*.

As previously explained, the underlining procedure is not a physically correct reproduction of the user's viewing condition, however it can lead to high accuracy of space perception, and therefore to a general improvement in performance in terms of realistic impression of depth. In fact, it is expected to compensate for visual misalignments in the rendered scene.

The viewpoint elevation and viewpoint retreat calibration actions are simple to implement. They can be performed quickly at the beginning of each individual observation. This can be implemented, for example, similarly to the popular eye-tracking calibration procedure often occurring soon after wearing a VR headset equipped with user's eye-tracking.

A fully automatic calibration of the proposed technique can be realistically expected in future implementation. It will require knowledge of the observed scene-portions for which high spatial precision is desired. This in turn means identification of the no-displacement zone within the rendered image.

Alternatively, the no-displacement zone can be indicated by the user while or before operation, within a semi-automatic calibration approach. For example, it could be provided by the user through a relatively long fixation of the eye on the target area.

Knowing the specific distance from the observer at which we want to achieve greater spatial precision, or the specific 3D-space region, do not directly tell the system which region of the rendered-image the specific distance corresponds to. This is a challenge towards automating the process due to the numerous parameters involved in mapping the acquired image to the displayed outcome.

Specific applications, and therefore scenarios, can solve this mapping problem by exploiting possible landmarks or reference points. Machine learning (ML) techniques can also represent a useful approach to consider. Algorithms such as the Yolo, can help with automatically estimate distances to objects observed in the scene, and therefore to identify what image-region the set distance corresponds to.

We have not experimented with an automated process for true-dimensional visualization, nonetheless we believe this can be well developed with the technological tools today available. Future research most relevant as follow up of the proposed technique should certainly consists of investigating the above-mentioned challenge, or the specific implementation of an automate AI/ML based solution, to provide correct image-space coordinates. This is needed for a simple implementation of the viewpoint elevation and viewpoint retreat proposed calibration action.

Immersive visualization, photo-based VR and true-dimensional visualization, hold a demonstrated value and have great application potential in VR. We deem their adoption will certainly grow in near future developments in the XR sector.

References

[1] S. Livatino, M. Mohamed, G. Morana, P. Gainley, Y. Iqbal, T. H. Nguyen, K. Williams and A. Zocco, ‘Are We Ready for Take-Off ? Learning Cockpit Actions with VR Headsets’, in *Extended Reality*, vol. 13445, L. T. De Paolis, P. Arpaia, and M. Sacco, Eds., in Lecture Notes in Computer Science, vol. 13445. , Cham: Springer International Publishing, 2022, pp. 147–153. doi: 10.1007/978-3-031-15546-8_13.

[2] S. Livatino, A. Zocco, Y. Iqbal, P. Gainley, G. Morana, and G. M. Farinella, ‘Effects of Head Rotation and Depth Enhancement in Virtual Reality User-Scene Interaction’, in *Extended Reality*, vol. 13445, L. T. De Paolis, P. Arpaia, and M. Sacco, Eds., in Lecture Notes in Computer Science, vol. 13445. , Cham: Springer International Publishing, 2022, pp. 139–146. doi: 10.1007/978-3-031-15546-8_12.

[3] S. Livatino, G. Morana, Y. Iqbal, M. Mohamed, S. Hwang, P. Gainley, T. H. Nguyen, K. Williams and A. Zocco, ‘Immersive Visualization in Pilot Training: from Cockpit Panels to Drone Navigation’, in *2022 IEEE International Conference on Metrology for Extended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE)*, Rome, Italy: IEEE, Oct. 2022, pp. 454–458. doi: 10.1109/MetroXRAINE54828.2022.9967530.

[4] S. Livatino, A. Zocco, Y. Iqbal, P. Gainley, G. Morana, and G. M. Farinella, ‘Virtual Reality User-Scene Interaction:Head-Rotation versus Joystick Movements’, in *2022 IEEE International Conference on Metrology for Extended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE)*, Oct. 2022, pp. 93–98. doi: 10.1109/MetroXRAINE54828.2022.9967622.

[5] A. Zocco, S. Livatino, P. Gainley, Y. Iqbal, and G. Morana, ‘The Immersion Advantage in Command & Control: from Desktop Monitors to VR Headsets’, in *2022 IEEE International Conference on Metrology for Extended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE)*, Rome, Italy: IEEE, Oct. 2022, pp. 449–453. doi: 10.1109/MetroXRAINE54828.2022.9967491.

[6] S. Livatino, A. Regalbuto, G. Morana, G. Signorello, G. Gallo, A. Torrisi, G. Padula, K. Pelc, A. Malizia and G. M. Farinella, ‘Photorealistic True-Dimensional Visualization of Remote Panoramic Views for VR Headsets’, *IEEE Access*, vol. 11, pp. 60305–60323, 2023, doi: 10.1109/ACCESS.2023.3285709.

[7] ‘360-degree camera RICOH THETA’. Accessed: Jul. 19, 2023. [Online]. Available: https://theta360.com/uk/

[8] ‘Insta360 ONE X - Possiedi il momento.’ Accessed: Jul. 19, 2023. [Online]. Available: https://www.insta360.com/it/product/insta360-onex

[9] ‘Vuze+ Camera | Vuze Camera’. Accessed: Jul. 19, 2023. [Online]. Available: https://vuze.camera/camera/vuze-plus-camera

[10] Insta360, ‘Camera 3D - 360° - Insta360 Pro 8K’. Accessed: Jul. 16, 2023. [Online]. Available: https://www.insta360.com/it/product/insta360-pro/

[11] 'Insta360 Titan - Videocamera professionale 360 VR 3D | Registrare in 11K'. Accessed: Jul. 19, 2023. [Online]. Available: https://www.insta360.com/it/product/insta360-titan/

[12] D. S. U. May 2, 2017 7:29 Pm Posted: Dec 5, and 2016 3:00 Pm, 'Oculus Touch Controllers Review', IGN. Accessed: Jul. 19, 2023. [Online]. Available: https://www.ign.com/articles/2016/12/05/oculus-touch-controllers-review

[13] 'Controller | VIVE United Kingdom'. Accessed: Jul. 19, 2023. [Online]. Available: https://www.vive.com/uk/accessory/controller/

[14] 'Leap Motion Controller 2 | Smaller, Lighter, Better | Ultraleap'. Accessed: Jul. 19, 2023. [Online]. Available: https://www.ultraleap.com/product/leap-2/

[15] 'Thumbs Up: Hand Tracking Available on Oculus Quest This Week | Meta Quest Blog'. Accessed: Jul. 19, 2023. [Online]. Available: https://www.oculus.com/blog/thumbs-up-hand-tracking-now-available-on-oculus-quest/

[16] T. Piumsomboon, G. Lee, R. W. Lindeman, and M. Billinghurst, 'Exploring natural eye-gaze-based interaction for immersive virtual reality', in *2017 IEEE Symposium on 3D User Interfaces (3DUI)*, Mar. 2017, pp. 36–39. doi: 10.1109/3DUI.2017.7893315.

[17] A. Regalbuto, S. Livatino, K. Edwards, and I. Mporas, 'Mobile VR headset usability evaluation of 2D and 3D panoramic views captured with different cameras', in *Interactive Collaborative Robotics: Second International Conference, ICR 2017, Hatfield, UK, September 12-16, 2017, Proceedings 2*, Springer, 2017, pp. 191–200.

[18] M. Brown and D. G. Lowe, 'Automatic Panoramic Image Stitching using Invariant Features', *Int. J. Comput. Vis.*, vol. 74, no. 1, pp. 59–73, Aug. 2007, doi: 10.1007/s11263-006-0002-3.

[19] R. F. Murray, K. Y. Patel, and E. S. Wiedenmann, 'Luminance calibration of virtual reality displays in Unity', *J. Vis.*, vol. 22, no. 13, p. 1, Dec. 2022, doi: 10.1167/jov.22.13.1.

[20] B. Lane, 'Stereoscopic displays', *Proc. SPIE*, pp. 20–32, 1982.

[21] C. Wheatstone, 'XVIII. Contributions to the physiology of vision.—Part the first. on some remarkable, and hitherto unobserved, phenomena of binocular vision', *Philosophical transactions of the Royal Society of London*, pp. 371–394, 1838.

[22] 'NVIDIA GeForce RTX 3080 Family', NVIDIA. Accessed: Jul. 19, 2023. [Online]. Available: https://www.nvidia.com/en-gb/geforce/graphics-cards/30-series/rtx-3080-3080ti/

[23] F. El Jamiy and R. Marsh, 'Survey on depth perception in head mounted displays: distance estimation in virtual reality, augmented reality, and mixed reality', *IET Image Process.*, vol. 13, no. 5, pp. 707–712, Apr. 2019, doi: 10.1049/iet-ipr.2018.5920.

[24] N. Neumann, 'US Air Force trials virtual reality for crew chief course', Airforce Technology. Accessed: Jul. 07, 2023. [Online]. Available: https://www.airforce-technology.com/features/us-air-force-trials-virtual-reality-for-crew-chief-course/

[25] A. Marquardt, C. Trepkowski, T. D. Eibich, J. Maiero, E. Kruijff, and J. Schöning, ‘Comparing Non-Visual and Visual Guidance Methods for Narrow Field of View Augmented Reality Displays’, *IEEE Trans. Vis. Comput. Graph.*, vol. 26, no. 12, pp. 3389–3401, Dec. 2020, doi: 10.1109/TVCG.2020.3023605.

[26] M. Pomplun, T. W. Garaas, and M. Carrasco, ‘The effects of task difficulty on visual search strategy in virtual 3D displays’, *J. Vis.*, vol. 13, no. 3, pp. 24–24, Aug. 2013, doi: 10.1167/13.3.24.

[27] E. D. Ragan, R. Kopper, P. Schuchardt, and D. A. Bowman, ‘Studying the Effects of Stereo, Head Tracking, and Field of Regard on a Small-Scale Spatial Judgment Task’, *IEEE Trans. Vis. Comput. Graph.*, vol. 19, no. 5, pp. 886–896, May 2013, doi: 10.1109/TVCG.2012.163.

[28] W. Ijsselsteijn, H. Ridder, J. Freeman, and S. Avons, ‘Presence: Concept, determinants and measurement’, *Proc. SPIE - Int. Soc. Opt. Eng.*, vol. 3959, Nov. 2000, doi: 10.1117/12.387188.

[29] M. Aksoy, C. Ufodiama, A. Bateson, S. Martin, and A. Asghar, ‘A comparative experimental study of visual brain event-related potentials to a working memory task: virtual reality head-mounted display versus a desktop computer screen’, *Exp. Brain Res.*, vol. 239, Oct. 2021, doi: 10.1007/s00221-021-06158-w.

[30] S. Sharples, S. Cobb, A. Moody, and J. R. Wilson, ‘Virtual reality induced symptoms and effects (VRISE): Comparison of head mounted display (HMD), desktop and projection display systems’, *Displays*, vol. 29, no. 2, pp. 58–69, Mar. 2008, doi: 10.1016/j.displa.2007.09.005.

[31] A. Bueckle, K. Buehling, P. Shih, and K. Borner, *Comparing Completion Time, Accuracy, and Satisfaction in Virtual Reality vs. Desktop Implementation of the Common Coordinate Framework Registration User Interface (CCF RUI)*. 2021.

[32] I. Horváth, ‘Behaviors and Capabilities of Generation CE Students in 3D VR’, in *2019 10th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, Oct. 2019, pp. 491–494. doi: 10.1109/CogInfoCom47531.2019.9089988.

[33] S. Deakin and F. Wilkinson, ‘The law of the labour market: Industrialization, employment, and legal evolution’, 2005.

[34] J. Bui, B. Li, E. Fung, M. Antonucci, K. D. Tran, S. Patel, S. Chung, D. Levi and R. Li, ‘Stereoscopic 3D video games boost depth perception’, *Invest. Ophthalmol. Vis. Sci.*, vol. 60, no. 9, p. 1797, Jul. 2019.

[35] A. Naceri, A. Moscatelli, and R. Chellali, ‘Depth discrimination of constant angular size stimuli in action space: role of accommodation and convergence cues’, *Front. Hum. Neurosci.*, vol. 9, 2015, Accessed: Jul. 07, 2023. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fnhum.2015.00511

[36] J. McIntire, P. Havig, and E. Geiselman, ‘Stereoscopic 3D displays and human performance: A comprehensive review’, *Displays*, vol. 35, pp. 18–26, Jan. 2014, doi: 10.1016/j.displa.2013.10.004.

[37] N. C. Nilsson, R. Nordahl, and S. Serafin, ‘Immersion Revisited: A review of existing definitions of immersion and their relation to different theories of presence’, *Hum. Technol.*, vol. 12, no. 2, pp. 108–134, Nov. 2016, doi: 10.17011/ht/urn.201611174652.

[38] S. Baceviciute, A. L. Cordoba, P. Wismer, T. V. Jensen, M. Klausen, and G. Makransky, ‘Investigating the value of immersive virtual reality tools for organizational training: An applied international study in the biotech industry’, *J. Comput. Assist. Learn.*, vol. 38, no. 2, pp. 470–487, 2022, doi: 10.1111/jcal.12630.

[39] E. B. Nash, G. W. Edwards, J. A. Thompson, and W. Barfield, ‘A Review of Presence and Performance in Virtual Environments’, *Int. J. Hum.-Comput. Interact.*, vol. 12, no. 1, pp. 1–41, May 2000, doi: 10.1207/S15327590IJHC1201_1.

[40] Z. Emami and T. Chau, ‘The effects of visual distractors on cognitive load in a motor imagery brain-computer interface’, *Behav. Brain Res.*, vol. 378, p. 112240, Jan. 2020, doi: 10.1016/j.bbr.2019.112240.

[41] B. Olk, A. Dinu, D. J. Zielinski, and R. Kopper, ‘Measuring visual search and distraction in immersive virtual reality’, *R. Soc. Open Sci.*, vol. 5, no. 5, p. 172331, May 2018, doi: 10.1098/rsos.172331.

[42] Z. (Jing) Xu, A. Lleras, Y. Shao, and S. Buetti, ‘Distractor–distractor interactions in visual search for oriented targets explain the increased difficulty observed in nonlinearly separable conditions.’, *J. Exp. Psychol. Hum. Percept. Perform.*, vol. 47, no. 9, pp. 1274–1297, Sep. 2021, doi: 10.1037/xhp0000941.

[43] A. Lleras, Z. Wang, A. Madison, and S. Buetti, ‘Predicting Search Performance in Heterogeneous Scenes: Quantifying the Impact of Homogeneity Effects in Efficient Search’, *Collabra Psychol.*, vol. 5, no. 1, p. 2, Jan. 2019, doi: 10.1525/collabra.151.

[44] W. T. Walters and J. Walton, ‘Efficacy of Virtual Reality Training for Pilots: A Review of Links between User Presence, Search Task Performance, and Collaboration within Virtual Reality’, *Proc. Hum. Factors Ergon. Soc. Annu. Meet.*, vol. 65, no. 1, pp. 919–922, Sep. 2021, doi: 10.1177/1071181321651347.

[45] G. Makransky and G. B. Petersen, ‘The Cognitive Affective Model of Immersive Learning (CAMIL): a Theoretical Research-Based Model of Learning in Immersive Virtual Reality’, *Educ. Psychol. Rev.*, vol. 33, no. 3, pp. 937–958, Sep. 2021, doi: 10.1007/s10648-020-09586-2.

[46] E. Krokos, C. Plaisant, and A. Varshney, ‘Virtual memory palaces: immersion aids recall’, *Virtual Real.*, vol. 23, no. 1, pp. 1–15, Mar. 2019, doi: 10.1007/s10055-018-0346-3.

[47] N. A. Stanton, P. M. Salmon, and G. H. Walker, ‘Let the Reader Decide: A Paradigm Shift for Situation Awareness in Sociotechnical Systems’, *J. Cogn. Eng. Decis. Mak.*, vol. 9, no. 1, pp. 44–50, Mar. 2015, doi: 10.1177/1555343414552297.

[48] M. Johnson-Glenberg, H. Bartolomea, and E. Kalina, ‘Platform is not destiny: Embodied learning effects comparing 2D desktop to 3D virtual reality STEM experiences’, *J. Comput. Assist. Learn.*, vol. 37, Oct. 2021, doi: 10.1111/jcal.12567.

[49] S. Livatino and C. Hochleitner, ‘Simple Guidelines for Testing VR Applications’, 2008. doi: 10.5772/5925.

[50] T. Schubert, ‘The sense of presence in virtual environments: A three-component scale measuring spatial presence, involvement, and realness’, *Z. Für Medien.*, vol. 15, pp. 69–71, Apr. 2003, doi: 10.1026//1617-6383.15.2.69.

[51] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal, ‘Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness’, *Int. J. Aviat. Psychol.*, vol. 3, no. 3, pp. 203–220, Jul. 1993, doi: 10.1207/s15327108ijap0303_3.

[52] S. A. Balk, M. A. Bertola, and V. W. Inman, ‘Simulator Sickness Questionnaire: Twenty Years Later’, in *Proceedings of the 7th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design : driving assessment 2013*, Bolton Landing, New York, USA>: University of Iowa, 2013, pp. 257–263. doi: 10.17077/drivingassessment.1498.

[53] J. Brooke, ‘SUS: a retrospective’, *J. Usability Stud.*, vol. 8, pp. 29–40, Jan. 2013.

[54] M. R. Endsley, ‘A Systematic Review and Meta-Analysis of Direct Objective Measures of Situation Awareness: A Comparison of SAGAT and SPAM’, *Hum. Factors*, vol. 63, no. 1, pp. 124–150, Feb. 2021, doi: 10.1177/0018720819875376.

[55] M. Ali and R. E. Cardona-Rivera, ‘Comparing Gamepad and Naturally-mapped Controller Effects on Perceived Virtual Reality Experiences’, in *ACM Symposium on Applied Perception 2020*, in SAP ’20. New York, NY, USA: Association for Computing Machinery, Sep. 2020, pp. 1–10. doi: 10.1145/3385955.3407923.

[56] J. Seibert and D. Shafer, ‘Control mapping in virtual reality: effects on spatial presence and controller naturalness’, *Virtual Real.*, vol. 22, pp. 1–10, Mar. 2018, doi: 10.1007/s10055-017-0316-1.

[57] M. Ebnali, R. Lamb, and R. Fathi, ‘Familiarization tours for first-time users of highly automated cars: Comparing the effects of virtual environments with different levels of interaction fidelity’, *ArXiv*, Feb. 2020, Accessed: Jul. 08, 2023. [Online]. Available: https://www.semanticscholar.org/paper/Familiarization-tours-for-first-time-users-of-cars%3A-Ebnali-Lamb/124281cf1b52add62b9244606c5ab79f2395f2a0

[58] E. D. Ragan, S. Scerbo, F. Bacim, and D. A. Bowman, ‘Amplified Head Rotation in Virtual Reality and the Effects on 3D Search, Training Transfer, and Spatial Orientation’, *IEEE Trans. Vis. Comput. Graph.*, vol. 23, no. 8, pp. 1880–1895, Aug. 2017, doi: 10.1109/TVCG.2016.2601607.

[59] S. P. Sargunam, K. R. Moghadam, M. Suhail, and E. D. Ragan, ‘Guided head rotation and amplified head rotation: Evaluating semi-natural travel and viewing techniques in virtual reality’, *2017 IEEE Virtual Real. VR*, pp. 19–28, 2017, doi: 10.1109/VR.2017.7892227.

[60] C. Christou, A. Tzanavari, K. Herakleous, and C. Poullis, ‘Navigation in virtual reality: Comparison of gaze-directed and pointing motion control’, in *2016 18th Mediterranean*

Electrotechnical Conference (MELECON), Apr. 2016, pp. 1–6. doi: 10.1109/MELCON.2016.7495413.

[61] R. Bayramova, I. Valori, P. E. McKenna-Plumley, C. Z. Callegher, and T. Farroni, ‘The role of vision and proprioception in self-motion encoding: An immersive virtual reality study’, *Atten. Percept. Psychophys.*, vol. 83, no. 7, pp. 2865–2878, Oct. 2021, doi: 10.3758/s13414-021-02344-8.

[62] E. D. Ragan, D. A. Bowman, R. Kopper, C. Stinson, S. Scerbo, and R. P. McMahan, ‘Effects of Field of View and Visual Complexity on Virtual Reality Training Effectiveness for a Visual Scanning Task’, *IEEE Trans. Vis. Comput. Graph.*, vol. 21, no. 7, pp. 794–807, Jul. 2015, doi: 10.1109/TVCG.2015.2403312.

[63] A. Zocco, M. D. Zocco, A. Greco, S. Livatino, and L. T. De Paolis, ‘Touchless Interaction for Command and Control in Military Operations’, in *Augmented and Virtual Reality*, L. T. De Paolis and A. Mongelli, Eds., in Lecture Notes in Computer Science. Cham: Springer International Publishing, 2015, pp. 432–445. doi: 10.1007/978-3-319-22888-4_32.

[64] J. Mateos-Garcia, K. Stathoulopoulos, and N. Thomas, ‘The immersive economy in the UK: Measuring the growth of virtual, augmented and mixed reality technologies’, nesta. Accessed: Jul. 08, 2023. [Online]. Available: https://www.nesta.org.uk/report/immersive-economy-uk/

[65] J. Ratcliffe and L. Tokarchuk, ‘Sensorimotor learning in immersive virtual reality: a scoping literature review’, in *2021 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*, Taichung, Taiwan: IEEE, Nov. 2021, pp. 276–286. doi: 10.1109/AIVR52153.2021.00061.

[66] S. A. Penumudi, V. A. Kuppam, J. H. Kim, and J. Hwang, ‘The effects of target location on musculoskeletal load, task performance, and subjective discomfort during virtual reality interactions’, *Appl. Ergon.*, vol. 84, p. 103010, Apr. 2020, doi: 10.1016/j.apergo.2019.103010.

[67] H. Leibowitz and D. Moore, ‘Role of Changes in Accommodation and Convergence in the Perception of Size*’, *JOSA*, vol. 56, no. 8, pp. 1120–1123, Aug. 1966, doi: 10.1364/JOSA.56.001120.

[68] R. S. Allison, B. J. Gillam, and E. Vecellio, ‘Binocular depth discrimination and estimation beyond interaction space’, *J. Vis.*, vol. 9, no. 1, p. 10, Jan. 2009, doi: 10.1167/9.1.10.

[69] M. Lamb, M. Brundin, E. Perez Luque, and E. Billing, ‘Eye-Tracking Beyond Peripersonal Space in Virtual Reality: Validation and Best Practices’, *Front. Virtual Real.*, vol. 3, 2022, Accessed: Jul. 11, 2023. [Online]. Available: https://www.frontiersin.org/articles/10.3389/frvir.2022.864653

[70] A. Zocco, S. Livatino, and L. T. De Paolis, ‘Stereoscopic-3D Vision to Improve Situational Awareness in Military Operations’, in *Augmented and Virtual Reality*, L. T. De Paolis and A. Mongelli, Eds., in Lecture Notes in Computer Science. Cham:

Springer International Publishing, 2014, pp. 351–362. doi: 10.1007/978-3-319-13969-2_26.

[71] K. Rahimi, C. Banigan, and E. D. Ragan, ‘Scene Transitions and Teleportation in Virtual Reality and the Implications for Spatial Awareness and Sickness’, *IEEE Trans. Vis. Comput. Graph.*, vol. 26, no. 6, pp. 2273–2287, Jun. 2020, doi: 10.1109/TVCG.2018.2884468.

[72] L. Rebenitsch, ‘Managing cybersickness in virtual reality’, *XRDS Crossroads ACM Mag. Stud.*, vol. 22, no. 1, pp. 46–51, Nov. 2015, doi: 10.1145/2810054.

[73] J. F. Golding, ‘Motion sickness susceptibility questionnaire revised and its relationship to other forms of sickness’, *Brain Res. Bull.*, vol. 47, no. 5, pp. 507–516, Nov. 1998, doi: 10.1016/S0361-9230(98)00091-4.

[74] H. Rao, R. Khanna, D. Zielinski, Y. Lu, J. Clements, N. Potter, M. Sommer, R. Kopper and L. Appelbaum, ‘Sensorimotor Learning during a Marksmanship Task in Immersive Virtual Reality’, *Front. Psychol.*, vol. 9, Feb. 2018, doi: 10.3389/fpsyg.2018.00058.

[75] ELT Group, ‘LOKI Command & Control System by Elettronica’. Accessed: Jul. 08, 2023. [Online]. Available: https://eltgroup.net/

[76] S. Kasahara and J. Rekimoto, ‘JackIn head: immersive visual telepresence system with omnidirectional wearable camera for remote collaboration’, in *Proceedings of the 21st ACM Symposium on Virtual Reality Software and Technology*, Beijing China: ACM, Nov. 2015, pp. 217–225. doi: 10.1145/2821592.2821608.

[77] J. A. Stevens and J. P. Kincaid, ‘The Relationship between Presence and Performance in Virtual Simulation Training’, *Open J. Model. Simul.*, vol. 03, no. 02, pp. 41–48, 2015, doi: 10.4236/ojmsi.2015.32005.

[78] M. Kraus, N. Weiler, D. A. Keim, A. Diehl, and B. Bach, ‘Visualization in the VR-Canvas: How much Reality is Good for Immersive Analytics in Virtual Reality?’.

[79] D. Patton, ‘How Real Is Good Enough? Assessing Realism of Presence in Simulations and Its Effects on Decision Making’, in *Foundations of Augmented Cognition. Advancing Human Performance and Decision-Making through Adaptive Systems*, vol. 8534, D. D. Schmorrow and C. M. Fidopiastis, Eds., in Lecture Notes in Computer Science, vol. 8534. , Cham: Springer International Publishing, 2014, pp. 245–256. doi: 10.1007/978-3-319-07527-3_23.

[80] B. G. Witmer and M. J. Singer, ‘Measuring Presence in Virtual Environments: A Presence Questionnaire’, *Presence Teleoperators Virtual Environ.*, vol. 7, no. 3, pp. 225–240, Jun. 1998, doi: 10.1162/105474698565686.

[81] R. B. Welch, T. T. Blackmon, A. Liu, B. A. Mellers, and L. W. Stark, ‘The Effects of Pictorial Realism, Delay of Visual Feedback, and Observer Interactivity on the Subjective Sense of Presence’, *Presence Teleoperators Virtual Environ.*, vol. 5, no. 3, pp. 263–273, Jan. 1996, doi: 10.1162/pres.1996.5.3.263.

[82] K. Jaalama, N. Fagerholm, A. Julin, J.-P. Virtanen, M. Maksimainen, and H. Hyyppä, ‘Sense of presence and sense of place in perceiving a 3D geovisualization for

communication in urban planning – Differences introduced by prior familiarity with the place', *Landsc. Urban Plan.*, vol. 207, p. 103996, Mar. 2021, doi: 10.1016/j.landurbplan.2020.103996.

[83] J. Diemer, G. W. Alpers, H. M. Peperkorn, Y. Shiban, and A. MÃ¼hlberger, 'The impact of perception and presence on emotional reactions: a review of research in virtual reality', *Front. Psychol.*, vol. 6, Jan. 2015, doi: 10.3389/fpsyg.2015.00026.

[84] M. C. Juan and D. Pérez, 'Comparison of the Levels of Presence and Anxiety in an Acrophobic Environment Viewed via HMD or CAVE', *Presence Teleoperators Virtual Environ.*, vol. 18, no. 3, pp. 232–248, Jun. 2009, doi: 10.1162/pres.18.3.232.

[85] S. Livatino, L. T. De Paolis, M. D'Agostino, A. Zocco, A. Agrimi, A. De Santis, L. V. Bruno and M. Lapresa, 'Stereoscopic Visualization and 3-D Technologies in Medical Endoscopic Teleoperation', *IEEE Trans. Ind. Electron.*, vol. 62, no. 1, pp. 525–535, Jan. 2015, doi: 10.1109/TIE.2014.2334675.

[86] D. A. Bowman and R. P. McMahan, 'Virtual Reality: How Much Immersion Is Enough?', *Computer*, vol. 40, no. 7, pp. 36–43, Jul. 2007, doi: 10.1109/MC.2007.257.

[87] A. Kuijsters, W. A. Ijsselsteijn, M. T. M. Lambooij, and I. E. J. Heynderickx, 'Influence of chroma variations on naturalness and image quality of stereoscopic images', presented at the IS&T/SPIE Electronic Imaging, B. E. Rogowitz and T. N. Pappas, Eds., San Jose, CA, Feb. 2009, p. 72401E. doi: 10.1117/12.817749.

[88] W. A. IJsselsteijn, H. De Ridder, and J. Vliegen, 'Subjective evaluation of stereoscopic images: effects of camera parameters and display duration', *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, no. 2, pp. 225–233, Mar. 2000, doi: 10.1109/76.825722.

[89] F. E. Jamiy and R. Marsh, 'Distance Estimation In Virtual Reality And Augmented Reality: A Survey', in *2019 IEEE International Conference on Electro Information Technology (EIT)*, May 2019, pp. 063–068. doi: 10.1109/EIT.2019.8834182.

[90] P. Seuntiens, L. Meesters, and W. Ijsselsteijn, 'Perceived quality of compressed stereoscopic images: Effects of symmetric and asymmetric JPEG coding and camera separation', *ACM Trans. Appl. Percept.*, vol. 3, no. 2, pp. 95–109, Apr. 2006, doi: 10.1145/1141897.1141899.

[91] V. Juřík, L. Herman, D. Snopková, A. J. Galang, Z. Stachoň, J. Chmelík, P. Kubíček and Č. Šašinka, 'The 3D hype: Evaluating the potential of real 3D visualization in geo-related applications', *PLOS ONE*, vol. 15, no. 5, p. e0233353, May 2020, doi: 10.1371/journal.pone.0233353.

[92] T. J. W. M. Janssen and F. J. J. Blommaert, 'A computational approach to image quality', *Displays*, vol. 21, no. 4, pp. 129–142, Oct. 2000, doi: 10.1016/S0141-9382(00)00056-1.

[93] R. G. Kaptein, A. Kuijsters, M. T. M. Lambooij, W. A. IJsselsteijn, and I. Heynderickx, 'Performance evaluation of 3D-TV systems', in *Image Quality and System Performance V*, SPIE, Jan. 2008, pp. 443–453. doi: 10.1117/12.770082.

[94] J. A. Ferwerda, ‘Three varieties of realism in computer graphics’, presented at the Electronic Imaging 2003, B. E. Rogowitz and T. N. Pappas, Eds., Santa Clara, CA, Jun. 2003, p. 290. doi: 10.1117/12.473899.

[95] M. A. Hagen, *Varieties of realism : geometries of representational art*. Cambridge ; New York : Cambridge University Press, 1986. Accessed: Jul. 09, 2023. [Online]. Available: http://archive.org/details/varietiesofreali0000hage

[96] M. S. Banks, D. M. Hoffman, J. Kim, and G. Wetzstein, ‘3D Displays’, *Annu. Rev. Vis. Sci.*, vol. 2, no. 1, pp. 397–435, Oct. 2016, doi: 10.1146/annurev-vision-082114-035800.

[97] A. Tiiro, ‘Effect of visual realism on cybersickness in virtual reality’, 2018. Accessed: Jul. 09, 2023. [Online]. Available: https://www.semanticscholar.org/paper/Effect-of-visual-realism-on-cybersickness-in-Tiiro/a4173881295298565c398418d39a8f8cf175b007

[98] P. J. Pardo, M. I. Suero, and Á. L. Pérez, ‘Correlation between perception of color, shadows, and surface textures and the realism of a scene in virtual reality’, *J. Opt. Soc. Am. A*, vol. 35, no. 4, p. B130, Apr. 2018, doi: 10.1364/JOSAA.35.00B130.

[99] M. Slater, M. Usoh, and Y. Chrysanthou, ‘The Influence of Dynamic Shadows on Presence in Immersive Virtual Environments’, in *Virtual Environments ’95*, M. Göbel, Ed., in Eurographics. Vienna: Springer, 1995, pp. 8–21. doi: 10.1007/978-3-7091-9433-1_2.

[100] M. Slater, M. Usoh, and A. Steed, ‘Depth of Presence in Virtual Environments’, *Presence Teleoperators Virtual Environ.*, vol. 3, no. 2, pp. 130–144, Jan. 1994, doi: 10.1162/pres.1994.3.2.130.

[101] V. Palad, ‘The Difference Between Clarity, Sharpness, and Contrast Sliders • Pixels and Wanderlust’. Accessed: Jul. 10, 2023. [Online]. Available: https://pixelsandwanderlust.com/the-difference-between-clarity-sharpness-and-contrast-sliders/

[102] K. Gu, G. Zhai, W. Lin, X. Yang, and W. Zhang, ‘No-Reference Image Sharpness Assessment in Autoregressive Parameter Space’, *IEEE Trans. Image Process.*, vol. 24, no. 10, pp. 3218–3231, Oct. 2015, doi: 10.1109/TIP.2015.2439035.

[103] A. F. R. Guarda, N. M. M. Rodrigues, and F. Pereira, ‘Deep Learning-Based Point Cloud Geometry Coding: RD Control Through Implicit and Explicit Quantization’, in *2020 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*, Jul. 2020, pp. 1–6. doi: 10.1109/ICMEW46912.2020.9106022.

[104] T. Schubert, F. Friedmann, and H. Regenbrecht, ‘The Experience of Presence: Factor Analytic Insights’, *Presence Teleoperators Virtual Environ.*, vol. 10, no. 3, pp. 266–281, Jun. 2001, doi: 10.1162/105474601300343603.

[105] V. Interrante, B. Ries, J. Lindquist, and L. Anderson, ‘Elucidating Factors that can Facilitate Veridical Spatial Perception in Immersive Virtual Environments’.

[106] W. IJsselsteijn, H. D. Ridder, J. Freeman, S. E. Avons, and D. Bouwhuis, ‘Effects of Stereoscopic Presentation, Image Motion, and Screen Size on Subjective and Objective Corroborative Measures of Presence’, *Presence Teleoperators Virtual Environ.*, vol. 10, no. 3, pp. 298–311, Jun. 2001, doi: 10.1162/105474601300343621.

[107] Y. Ling, H. T. Nefs, W.-P. Brinkman, C. Qu, and I. Heynderickx, ‘The Effect of Perspective on Presence and Space Perception’, *PLoS ONE*, vol. 8, no. 11, p. e78513, Nov. 2013, doi: 10.1371/journal.pone.0078513.

[108] J. S. Hvass, O. Larsen, K. B. Vendelbo, N. C. Nilsson, R. Nordahl, and S. Serafin, ‘The effect of geometric realism on presence in a virtual reality game’, in *2017 IEEE Virtual Reality (VR)*, Los Angeles, CA, USA: IEEE, 2017, pp. 339–340. doi: 10.1109/VR.2017.7892315.

[109] M. Huang and N. Alessi, ‘Presence as an emotional experience: (705302011-011)’. 1999. doi: 10.1037/e705302011-011.

[110] R. M. Baños, E. Etchemendy, D. Castilla, A. García-Palacios, S. Quero, and C. Botella, ‘Positive mood induction procedures for virtual environments designed for elderly people’, *Interact. Comput.*, vol. 24, no. 3, pp. 131–138, May 2012, doi: 10.1016/j.intcom.2012.04.002.

[111] F. J. Seagull and D. M. Rooney, ‘Filling a void: Developing a standard subjective assessment tool for surgical simulation through focused review of current practices’, *Surgery*, vol. 156, no. 3, pp. 718–722, Sep. 2014, doi: 10.1016/j.surg.2014.04.048.

[112] J. Häkkinen, T. Kawai, J. Takatalo, T. Leisti, J. Radun, A. Hirsaho and G. Nyman, ‘Measuring stereoscopic image quality experience with interpretation based quality methodology’, presented at the Electronic Imaging 2008, S. P. Farnand and F. Gaykema, Eds., San Jose, CA, Jan. 2008, p. 68081B. doi: 10.1117/12.760935.

[113] E. Klein, J. E. Swan, G. S. Schmidt, M. A. Livingston, and O. G. Staadt, ‘Measurement Protocols for Medium-Field Distance Perception in Large-Screen Immersive Displays’, in *2009 IEEE Virtual Reality Conference*, Lafayette, LA: IEEE, Mar. 2009, pp. 107–113. doi: 10.1109/VR.2009.4811007.

[114] I. T. Feldstein, F. M. Kölsch, and R. Konrad, ‘Egocentric Distance Perception: A Comparative Study Investigating Differences Between Real and Virtual Environments’, *Perception*, vol. 49, no. 9, pp. 940–967, Sep. 2020, doi: 10.1177/0301006620951997.

[115] E. Brivio, S. Serino, E. Negro Cousa, A. Zini, G. Riva, and G. De Leo, ‘Virtual reality and 360° panorama technology: a media comparison to study changes in sense of presence, anxiety, and positive emotions’, *Virtual Real.*, vol. 25, no. 2, pp. 303–311, Jun. 2021, doi: 10.1007/s10055-020-00453-7.

[116] V. Interrante, B. Ries, and L. Anderson, ‘Distance Perception in Immersive Virtual Environments, Revisited’, in *IEEE Virtual Reality Conference (VR 2006)*, Mar. 2006, pp. 3–10. doi: 10.1109/VR.2006.52.

[117] W. B. Thompson, P. Willemsen, A. A. Gooch, S. H. Creem-Regehr, J. M. Loomis, and A. C. Beall, ‘Does the Quality of the Computer Graphics Matter when Judging

Distances in Visually Immersive Environments?', *Presence*, vol. 13, no. 5, pp. 560–571, Oct. 2004, doi: 10.1162/1054746042545292.

[118] R. Newell, R. Canessa, and T. Sharma, 'Visualizing Our Options for Coastal Places: Exploring Realistic Immersive Geovisualizations as Tools for Inclusive Approaches to Coastal Planning and Management', *Front. Mar. Sci.*, vol. 4, p. 290, Sep. 2017, doi: 10.3389/fmars.2017.00290.

[119] M. Davenport and D. Anderson, 'Getting From Sense of Place to Place-Based Management: An Interpretive Investigation of Place Meanings and Perceptions of Landscape Change', *Soc. Nat. Resour. - SOC Nat. RESOUR*, vol. 18, pp. 625–641, Aug. 2005, doi: 10.1080/08941920590959613.

[120] A. Julin, K. Jaalama, J.P. Virtanen, M. Maksimainen, M. Kurkela, J. Hyyppä and H. Hyyppä, 'Automated Multi-Sensor 3D Reconstruction for the Web', *ISPRS Int. J. Geo-Inf.*, vol. 8, no. 5, Art. no. 5, May 2019, doi: 10.3390/ijgi8050221.

[121] J. P. Virtanen, M. Kurkela, T. Turppa, M. T. Vaaja, A. Julin, A. Kukko, J. Hyyppä, M. Ahlavuo, J. E. Von Numers, H. Haggrén and H. Hyyppä, 'Depth camera indoor mapping for 3D virtual radio play', *Photogramm. Rec.*, vol. 33, no. 162, pp. 171–195, 2018, doi: 10.1111/phor.12239.

[122] T. Rhee, S. Thompson, D. Medeiros, R. Dos Anjos, and A. Chalmers, 'Augmented Virtual Teleportation for High-Fidelity Telecollaboration', *IEEE Trans. Vis. Comput. Graph.*, vol. 26, no. 5, pp. 1923–1933, May 2020, doi: 10.1109/TVCG.2020.2973065.

[123] P. Fuchs, 'Interfaces visuelles', éditeur de l'Encyclopédie des Techniques de l'ingénieur, 2003, vol. vol. TE 5906. Accessed: Jul. 10, 2023. [Online]. Available: https://mines-paristech.hal.science/hal-00785602

[124] J. A. Da Silva, 'Scales for Perceived Egocentric Distance in a Large Open Field: Comparison of Three Psychophysical Methods', *Am. J. Psychol.*, vol. 98, no. 1, pp. 119–144, 1985, doi: 10.2307/1422771.

[125] G. von Békésy, *Experiments in hearing*. in McGraw-Hill series in psychology. New York: McGraw-Hill, 1960.

[126] J. E. Cutting and P. M. Vishton, 'Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth'.

[127] P. Willemsen, A. A. Gooch, W. B. Thompson, and S. H. Creem-Regehr, 'Effects of Stereo Viewing Conditions on Distance Perception in Virtual Environments', *Presence Teleoperators Virtual Environ.*, vol. 17, no. 1, pp. 91–101, Feb. 2008, doi: 10.1162/pres.17.1.91.

[128] J. Knapp and J. Loomis, 'Visual Perception of Egocentric Distance in Real and Virtual Environments', in *Virtual Adapt. Environ.*, vol. 11, 2003, pp. 21–46. doi: 10.1201/9781410608888.pt1.

[129] P. B. Hibbard, A. E. Haines, and R. L. Hornsey, 'Magnitude, precision, and realism of depth perception in stereoscopic vision', *Cogn. Res. Princ. Implic.*, vol. 2, no. 1, p. 25, Dec. 2017, doi: 10.1186/s41235-017-0062-7.

[130] J. P. McIntire and K. K. Liggett, 'The (possible) utility of stereoscopic 3D displays for information visualization: The good, the bad, and the ugly', in *2014 IEEE VIS International Workshop on 3DVis (3DVis)*, Paris, France: IEEE, Nov. 2014, pp. 1–9. doi: 10.1109/3DVis.2014.7160093.

[131] E. Wilson, D. G. Hewett, B. C. Jolly, S. Janssens, and M. M. Beckmann, 'Is that realistic? The development of a realism assessment questionnaire and its application in appraising three simulators for a gynaecology procedure', *Adv. Simul.*, vol. 3, no. 1, p. 21, Dec. 2018, doi: 10.1186/s41077-018-0080-7.

[132] S. J. Hamstra, R. Brydges, R. Hatala, B. Zendejas, and D. A. Cook, 'Reconsidering Fidelity in Simulation-Based Training':, *Acad. Med.*, vol. 89, no. 3, pp. 387–392, Mar. 2014, doi: 10.1097/ACM.0000000000000130.

[133] D. E. Brackney and K. Priode, 'Back to Reality: The Use of the Presence Questionnaire for Measurement of Fidelity in Simulation', *J. Nurs. Meas.*, vol. 25, no. 2, pp. 66–73, Aug. 2017, doi: 10.1891/1061-3749.25.2.E66.

[134] L. Lioce, C. H. Meakim, M. K. Fey, J. V. Chmil, B. Mariani, and G. Alinier, 'Standards of best practice: Simulation design standard IX Clinical Simulation in Nursing, vol. 11, no. 6, pp. 309--315, 2015.', *Clin. Simul. Nurs.*, vol. 11, no. 3, pp. 309--315, 2015.

[135] A. Hill, M. S. Horswill, A. M. Plooy, M. O. Watson, R. Karamatic, T. A. Basit, G. M. Wallis, S. Riek, R. Burgess-Limerick and D. G. Hewett, 'Assessing the realism of colonoscopy simulation: the development of an instrument and systematic comparison of 4 simulators', *Gastrointest. Endosc.*, vol. 75, no. 3, pp. 631-640.e3, Mar. 2012, doi: 10.1016/j.gie.2011.10.030.

[136] Igroup, 'igroup presence questionnaire (IPQ)'. Online. [Online]. Available: http://www.igroup.org/pq/ipq/index.php

[137] D. Kasik, J. J. Troy, S. R. Amorosi, M. O. Murray, and S. N. Swamy, 'Evaluating graphics displays for complex 3D models', *Comput. Graph. Appl. IEEE*, vol. 22, pp. 56–64, Jun. 2002, doi: 10.1109/MCG.2002.999788.

[138] J. Rubin, 'Handbook of Usability Testing: How to Plan, Design and Conduct Effective Tests', Jan. 2008.

[139] W. L. in R.-B. U. Experience, 'Usability Engineering : Book by Jakob Nielsen', Nielsen Norman Group. Accessed: Jul. 10, 2023. [Online]. Available: https://www.nngroup.com/books/usability-engineering/

[140] J. F. C. de Winter and D. Dodou, 'Five-Point Likert Items: t test versus Mann-Whitney-Wilcoxon (Addendum added October 2012)', doi: 10.7275/BJ1P-TS64.

[141] Z. Gao, A. Hwang, G. Zhai, and E. Peli, 'Correcting geometric distortions in stereoscopic 3D imaging', *PLOS ONE*, vol. 13, no. 10, p. e0205032, Oct. 2018, doi: 10.1371/journal.pone.0205032.

[142] Cyclopean Isles, 'Cyclopean masonry', *Wikipedia*. Jul. 10, 2023. Accessed: Jul. 10, 2023. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Cyclopean_masonry&oldid=1164720894

[143] Monello Cave - Italy, ‘Grotta Monello’. Accessed: Jul. 10, 2023. [Online]. Available: https://www.unict.it/it/terza-missione/riserve-naturali/rni-grotta-monello

[144] S. Livatino and A. Regalbuto, ‘3D Tour - Lachea Islas and Monello Cave’. Accessed: Jul. 10, 2023. [Online]. Available:
https://cutgana.unict.it/sites/cutgana.unict.it/VirtualTours3D/

[145] T. Ni, D. A. Bowman, and J. Chen, ‘Increased Display Size and Resolution Improve Task Performance in Information-Rich Virtual Environments’.

[146] M. Slater, D.-P. Pertaub, and A. Steed, ‘Public speaking in virtual reality: facing an audience of avatars’, *IEEE Comput. Graph. Appl.*, vol. 19, no. 2, pp. 6–9, Mar. 1999, doi: 10.1109/38.749116.

[147] C. Hendrix and W. Barfield, ‘Presence within Virtual Environments as a Function of Visual Display Parameters’, *Presence Teleoperators Virtual Environ.*, vol. 5, no. 3, pp. 274–289, Aug. 1996, doi: 10.1162/pres.1996.5.3.274.

[148] S. Livatino, G. Mattiolo, C. Castello, and C. Randazzo, ‘The Role of Photo-Realism in Virtual Reality Games’.

[149] R. G. Eggleston, W. P. Janson, and K. A. Aldrich, ‘Virtual reality system effects on size-distance judgements in a virtual environment’, in *Proceedings of the IEEE 1996 Virtual Reality Annual International Symposium*, Mar. 1996, pp. 139–146. doi: 10.1109/VRAIS.1996.490521.

[150] F. Pallavicini, A. Pepe, A. Ferrari, G. Garcea, A. Zanacchi, and F. Mantovani, ‘What Is the Relationship Among Positive Emotions, Sense of Presence, and Ease of Interaction in Virtual Reality Systems? An On-Site Evaluation of a Commercial Virtual Experience’, *Presence Teleoperators Virtual Environ.*, vol. 27, no. 2, pp. 183–201, Feb. 2018, doi: 10.1162/pres_a_00325.

[151] A. K. Seth, K. Suzuki, and H. D. Critchley, ‘An Interoceptive Predictive Coding Model of Conscious Presence’, *Front. Psychol.*, vol. 2, 2012, doi: 10.3389/fpsyg.2011.00395.

[152] R. S. Renner, B. M. Velichkovsky, and J. R. Helmert, ‘The perception of egocentric distances in virtual environments - A review’, *ACM Comput. Surv.*, vol. 46, no. 2, p. 23:1-23:40, Dec. 2013, doi: 10.1145/2543581.2543590.

[153] L. Phillips, B. Ries, V. Interrante, M. Kaeding, and L. Anderson, ‘Distance perception in NPR immersive virtual environments, revisited’, in *Proceedings of the 6th Symposium on Applied Perception in Graphics and Visualization*, Chania, Crete Greece: ACM, Sep. 2009, pp. 11–14. doi: 10.1145/1620993.1620996.

[154] S. M. Moore and M. N. Geuss, ‘Familiarity with teammate’s attitudes improves team performance in virtual reality’, *PLOS ONE*, vol. 15, no. 10, p. e0241011, Oct. 2020, doi: 10.1371/journal.pone.0241011.

[155] R. Epstein, J. Higgins, K. Jablonski, and A. Feiler, ‘Visual Scene Processing in Familiar and Unfamiliar Environments’, *J. Neurophysiol.*, vol. 97, pp. 3670–83, Jun. 2007, doi: 10.1152/jn.00003.2007.

[156] A. Ahrens, K. D. Lund, M. Marschall, and T. Dau, ‘Sound source localization with varying amount of visual information in virtual reality’, *PLOS ONE*, vol. 14, no. 3, p. e0214603, Mar. 2019, doi: 10.1371/journal.pone.0214603.

[157] W. D. Hairston, M. T. Wallace, J. W. Vaughan, B. E. Stein, J. L. Norris, and J. A. Schirillo, ‘Visual Localization Ability Influences Cross-Modal Bias’, *J. Cogn. Neurosci.*, vol. 15, no. 1, pp. 20–29, Jan. 2003, doi: 10.1162/089892903321107792.

[158] B. Li, R. Zhang, A. Nordman, and S. A. Kuhl, ‘The effects of minification and display field of view on distance judgments in real and HMD-based environments’, in *Proceedings of the ACM SIGGRAPH Symposium on Applied Perception*, in SAP ’15. New York, NY, USA: Association for Computing Machinery, Sep. 2015, pp. 55–58. doi: 10.1145/2804408.2804427.

[159] B. Li, J. Walker, and S. A. Kuhl, ‘The Effects of Peripheral Vision and Light Stimulation on Distance Judgments Through HMDs’, *ACM Trans. Appl. Percept.*, vol. 15, no. 2, pp. 1–14, Apr. 2018, doi: 10.1145/3165286.

[160] F. Kellner, B. Bolte, G. Bruder, U. Rautenberg, F. Steinicke, M. Lappe and R. Koch, ‘Geometric Calibration of Head-Mounted Displays and its Effects on Distance Estimation’, *IEEE Trans. Vis. Comput. Graph.*, vol. 18, pp. 589–96, Apr. 2012, doi: 10.1109/TVCG.2012.45.

[161] S. Livatino, G. Muscato, and F. Privitera, ‘Stereo Viewing and Virtual Reality Technologies in Mobile Robot Teleguide’, *IEEE Trans. Robot.*, vol. 25, no. 6, pp. 1343–1355, Dec. 2009, doi: 10.1109/TRO.2009.2028765.

[162] S. Baek and C. Lee, ‘Depth perception estimation of various stereoscopic displays’, *Opt. Express*, vol. 24, pp. 23618–23634, Oct. 2016, doi: 10.1364/OE.24.023618.

[163] J. C. A. Read, ‘Stereo vision and strabismus’, *Eye Lond. Engl.*, vol. 29, no. 2, pp. 214–224, Feb. 2015, doi: 10.1038/eye.2014.279.

[164] G. Bruder, F. A. Sanz, A.-H. Olivier, and A. Lecuyer, ‘Distance estimation in large immersive projection systems, revisited’, in *2015 IEEE Virtual Reality (VR)*, Mar. 2015, pp. 27–32. doi: 10.1109/VR.2015.7223320.

[165] J. S. Lappin, A. L. Shelton, and J. J. Rieser, ‘Environmental context influences visually perceived distance’, *Percept. Psychophys.*, vol. 68, no. 4, pp. 571–581, May 2006, doi: 10.3758/bf03208759.

[166] R. S. Renner, B. M. Velichkovsky, and J. R. Helmert, ‘The perception of egocentric distances in virtual environments - A review’, *ACM Comput. Surv.*, vol. 46, no. 2, p. 23:1-23:40, Dec. 2013, doi: 10.1145/2543581.2543590.

[167] Insta360, ‘Software: Insta360 Pro Control Camera App’. Accessed: Jul. 16, 2023. [Online]. Available: https://www.insta360.com/it/download/insta360-pro

[168] ‘Unity Real-Time Development Platform | 3D, 2D, VR & AR Engine’. Accessed: Jul. 16, 2023. [Online]. Available: https://unity.com/

[169] S. H. Creem-Regehr, J. K. Stefanucci, and B. Bodenheimer, ‘Perceiving distance in virtual reality: theoretical insights from contemporary technologies’, *Philos. Trans. R.*

Soc. B Biol. Sci., vol. 378, no. 1869, p. 20210456, Jan. 2023, doi: 10.1098/rstb.2021.0456.

[170] T. Combe, J.-R. Chardonnet, F. Merienne, and J. Ovtcharova, ‘CAVE and HMD: distance perception comparative study’, *Virtual Real.*, Mar. 2023, doi: 10.1007/s10055-023-00787-y.

Figures

Tables